

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 1, 2003, 16:02:29 ; Search time 25 seconds

(without alignments)
3262.219 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLFTNPLSPETAMREPAP.....VOKRGYYELVNLQSLGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 segs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PC1_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PC10_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6721	100.0	1334	9	US-09-758-828-2
2	2489	37.0	1280	9	US-10-072-621-7
3	2489	37.0	1280	10	US-09-866-866A-2
4	2487.5	37.0	1272	10	US-09-769-097-4
5	2482	36.9	1280	9	US-10-044-671-2
6	2480	36.9	1280	10	US-09-866-866A-4
7	2475.5	36.8	1272	10	US-09-769-097-2
8	2475.5	36.8	1276	10	US-09-866-866A-8
9	2453.5	36.5	1276	10	US-09-866-866A-6
10	2282	34.0	1195	10	US-09-873-409-6
11	2258.5	33.6	1222	10	US-09-873-409-5
12	2160.5	32.1	1275	10	US-09-749-340-6
13	2067	30.8	1058	10	US-09-873-409-4
14	1940	28.9	1263	9	US-09-882-694-11
15	1690.5	25.2	812	10	US-09-873-409-2
16	1617	24.1	1394	9	US-10-101-388-3
17	1254	18.7	659	10	US-09-873-409-1
18	1001.5	14.9	1247	9	US-09-738-626-4/51
19	937	13.9	514	10	US-09-873-409-8

20	913.5	13.6	541	10	US-09-873-409-7	Sequence 7, Appl1
21	814.5	12.1	766	9	US-10-072-621-6	Sequence 6, Appl1
22	813.5	12.1	766	9	US-10-156-239-17	Sequence 17, Appl1
23	813.5	12.1	766	10	US-09-795-693-17	Sequence 17, Appl1
24	756	11.2	582	10	US-09-815-242-14006	Sequence 14006, A
25	749.5	11.2	582	10	US-09-815-242-10115	Sequence 10115, A
26	740.5	11.0	1464	9	US-10-012-896-1008	Sequence 1008, Ap
27	739.5	11.0	1261	9	US-10-012-896-1009	Sequence 538, App
28	739.5	11.0	1261	9	US-10-012-896-1009	Sequence 538, App
29	739.5	11.0	1261	9	US-09-895-793-538	Sequence 538, App
30	739.5	11.0	1261	9	US-09-895-814-538	Sequence 538, App
31	739.5	11.0	1261	10	US-09-758-143-538	Sequence 538, App
32	739.5	11.0	1261	10	US-09-780-669-538	Sequence 538, App
33	739.5	11.0	1261	10	US-09-822-827-538	Sequence 538, App
34	720	10.7	1382	9	US-10-087-782A-31	Sequence 31, Appl1
35	702	10.4	1360	9	US-09-934-421A-2	Sequence 2, Appl1
36	702	10.4	1360	9	US-10-162-012-34	Sequence 34, Appl1
37	691.5	10.3	603	10	US-09-815-242-5202	Sequence 5202, Ap
38	682	10.1	748	9	US-09-870-759-41	Sequence 41, Appl1
39	674.5	10.0	1437	9	US-10-154-452-2	Sequence 2, Appl1
40	674.5	10.0	1437	9	US-09-934-421A-6	Sequence 6, Appl1
41	674.5	10.0	1437	9	US-10-162-012-38	Sequence 38, Appl1
42	670	10.0	1228	9	US-10-012-896-537	Sequence 537, App
43	670	10.0	1228	9	US-09-895-793-537	Sequence 537, App
44	670	10.0	1228	9	US-09-895-814-537	Sequence 537, App
45	670	10.0	1228	10	US-09-759-143-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-09-758-828-2
Sequence 2, Application US/09758828
Patent No. US20020164772A1
GENERAL INFORMATION:
APPLICANT: Skatrud, Paul L.
de Waard, Maarten A.
Peery, Robert B.
Andrade, Alan C.
TITLE OF INVENTION: Multiple Drug Resistance Gene atrod of
Aspergillus nidulans
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/758, 828
FILING DATE: 11-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/996,545
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1334 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-758-828-2

Query Match 100.0%; Score 6721; DB 9; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPLTNPSPPTAMREPASTSTTEQASTPHADEKTLSDLSAPSSSTTAPADKHNRP 60
DB 1 MSPLTNPSPPTAMREPASTSTTEQASTPHADEKTLSDLSAPSSSTTAPADKHNRP 60
QY 61 KSSSNNAVSVNEVDALTAHPEDEKQVLTQLEEKVNIISFGAMRYATKDDILIMVS 120
DB 61 KSSSNNAVSVNEVDALTAHPEDEKQVLTQLEEKVNIISFGAMRYATKDDILIMVS 120
QY 121 TICATAAATSTFORIMLYOISYDEFDLTKNLYEVLIGEFVTVYVSTVGFITGEHA 180
DB 121 TICATAAATSTFORIMLYOISYDEFDLTKNLYEVLIGEFVTVYVSTVGFITGEHA 180
QY 181 TOKIREYVLESLRONIGYFDKLAGEVYTRITADTNLIQDISEKVGITLTALATPVA 240
DB 181 TOKIREYVLESLRONIGYFDKLAGEVYTRITADTNLIQDISEKVGITLTALATPVA 240
QY 241 FIATVYKWKIALICSTTVALVLMGGSGOFIIRKSKLDSYGAGTVAEEVISTIN 300
DB 241 FIATVYKWKIALICSTTVALVLMGGSGOFIIRKSKLDSYGAGTVAEEVISTIN 300
QY 301 ATAFGTODKLAQYEVHLEDEAKMGTKNQIYVGMIGAMFGLMNSYGLGFMMGSRFLVD 360
DB 301 ATAFGTODKLAQYEVHLEDEAKMGTKNQIYVGMIGAMFGLMNSYGLGFMMGSRFLVD 360
QY 361 GAVDYGDLITVLMALLIGSFSIGANVSPNAQATNAVAAAATFGTITDROSPDIPSNBCK 420
DB 361 GAVDYGDLITVLMALLIGSFSIGANVSPNAQATNAVAAAATFGTITDROSPDIPSNBCK 420
QY 421 TLDEHGEHLKNNVHIPTSPREVTVMEDVSLSPAGKTALVVGSGSKSVVGLVERF 480
DB 421 TLDEHGEHLKNNVHIPTSPREVTVMEDVSLSPAGKTALVVGSGSKSVVGLVERF 480
QY 481 YMPVGTVLDDGHDIKDLRLRQOISLVSOEPLYFGTITKNIIRHGLIGTKYENESD 540
DB 481 YMPVGTVLDDGHDIKDLRLRQOISLVSOEPLYFGTITKNIIRHGLIGTKYENESD 540
QY 541 KYRELIEENAKKANAHDFTALPEGETNVGQRFLLSGGQKORAIARAAYVSDPKILL 600
DB 541 KYRELIEENAKKANAHDFTALPEGETNVGQRFLLSGGQKORAIARAAYVSDPKILL 600
QY 601 DEATSALDTKSGVVOALERAEGRTTIVIAHRLSTIKTANIVVVGKTAEGTDE 660
DB 601 DEATSALDTKSGVVOALERAEGRTTIVIAHRLSTIKTANIVVVGKTAEGTDE 660
QY 661 LVDRGAVKRLVEAQRINEQKADALEDAEDLTNADIAKIKITASASDLDGKPTTID 720
DB 661 LVDRGAVKRLVEAQRINEQKADALEDAEDLTNADIAKIKITASASDLDGKPTTID 720
QY 721 RTGTHKSVSAILSKRPETTPKYSMLTLKFVASFNRPEIPYMLIGLVFSLAGGGPT 780
DB 721 RTGTHKSVSAILSKRPETTPKYSMLTLKFVASFNRPEIPYMLIGLVFSLAGGGPT 780
QY 781 QAVYAKAISTSLPESQSKLRHADFMSLMEFVVGITOTOSTNGAAAFVCSERLIR 840
DB 781 QAVYAKAISTSLPESQSKLRHADFMSLMEFVVGITOTOSTNGAAAFVCSERLIR 840
QY 841 RARSTAFPTIIRODIAFDEKENSSTGALTSLSTETKHLGSGVGTTLGTLTSTLGA 900
DB 841 RARSTAFPTIIRODIAFDEKENSSTGALTSLSTETKHLGSGVGTTLGTLTSTLGA 900
QY 901 IIALAIGMKIALVCISVVPVLLACGFRTYMLAQFQSRSKLAVGSAFACATSSIRT 960
DB 901 IIALAIGMKIALVCISVVPVLLACGFRTYMLAQFQSRSKLAVGSAFACATSSIRT 960

QY 961 VASLTRERDWEIYHAQIDAOGRISLISYBSLLYASQAALVFECVALGFMYGTLGH 1020
DB 961 VASLTRERDWEIYHAQIDAOGRISLISYBSLLYASQAALVFECVALGFMYGTLGH 1020
QY 1021 HEYDIFREFVCESEILFGAAGTAVFSPADMGKAKNAAEFRRLFDKRPQIDNNEGE 1080
DB 1021 HEYDIFREFVCESEILFGAAGTAVFSPADMGKAKNAAEFRRLFDKRPQIDNNEGE 1080
QY 1081 KLEIVGEIEFRNVHRRPTREPQVLRGLDLYKPGQYVALVPSGCGKSTTALLERF 1140
DB 1081 KLEIVGEIEFRNVHRRPTREPQVLRGLDLYKPGQYVALVPSGCGKSTTALLERF 1140
QY 1141 YDAINGSLVDKDISKLNINSRSLISYSEPLLYOGTIENTLIGVEDVDEEPLI 1200
DB 1141 YDAINGSLVDKDISKLNINSRSLISYSEPLLYOGTIENTLIGVEDVDEEPLI 1200
QY 1201 KACKDANIYDFIMSLPEGRNTVVGSKGMLSGGQKORAVIARALLRDKILLDATSAL 1260
DB 1201 KACKDANIYDFIMSLPEGRNTVVGSKGMLSGGQKORAVIARALLRDKILLDATSAL 1260
QY 1261 DSESEKVOALDAARAGRTTIAVAHRUSTQKADVIYFDGKIVESGTHSELVQKGR 1320
DB 1261 DSESEKVOALDAARAGRTTIAVAHRUSTQKADVIYFDGKIVESGTHSELVQKGR 1320
QY 1321 YVELVNIOSLGKH 1334
DB 1321 YVELVNIOSLGKH 1334

RESULT 2 US-10-072-621-7

; Sequence 7, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; FILE REFERENCE: 100103.402
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-621-7

Query Match 37.0%; Score 2489; DB 9; Length 1280;
Best Local Similarity 42.1%; Pred. No. 5.9e-177;
Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

QY 94 EIKKNISFEGIMRYATKMDILIMYSTICATAASTFORIMLY----- 137
DB 28 KKKPTVSVFSMFRYSNMWDKLYMVGTLAIIHGAGLPLMLVFGEMTDIFANAGNED 87
QY 138 -----QISYDEFY-----DELTKNLYEVLIGEFVTVYVSTVGFITGEHATOKI 184
DB 88 LMSNTNNSDINDITGFNNLEEDMTIRAYIYSGIAGVVAIYVSWCCLAAAGQIHKI 147
QY 185 REYVLESLRONIGYFDKLAGEVYTRITADTNLIQDISEKVGITLTALATPVAFTA 244
DB 148 RKQFHAIMROBIGFVDVDELMTRLDVSKEINGIGDKIGMFPQSMATFFGFIYG 207
QY 245 YKWKIALICSTTVALVLMGGSGOFIIRKSKLDSYGAGTVAEEVISTINATAF 304
DB 208 FTRGKMLTLVLIATSPVGLSAVMAKILISFTDELLAYAKAGAAVEAIAITVIAF 267
QY 305 GTQDILAKQYEVHLEDEAKMGTKNQIYVGMIGAMFGLMNSYGLGFMMGSRFLVDGAVD 364
DB 268 GGQKLEIRYNKNLEAKRIGIKAITANISIGAAFLIITASYALAFWGTITLVLSGYS 327

QY	365	VGDITVLMIALLISFISGVNVPAAOFETNVAANAATIFGIDIROSPDLIDYSNBGRKIDH	424
Db	328	IGOVLTPEFVSLIAFVSQASPSIEFPAANRGAAYETIEKIIDNKSIDSTSKSGAKPDN	387
QY	425	FECHIELRNKHIYPSRPEVTIMEDVSI,SMRPAKTTALVPSRSGSKSTVGLVERFPMV	484
Db	388	IKGNLEFRNRYFVSPSRKEVKILKGLMKILOSQOTVALVNSGGCKSTVQLOMRLYDPT	447
QY	485	RGTVLLDGHODKIDNLEWMLNQOISLVSOEPUVLGCTTYKNIIRGLSTIKENESDKVRE	544
Db	448	EGMVSVDGODIRTNVFFLRITIGVVSQOEPUVLATTAENIRYG----RENTYMD-----	499
QY	545	LIEANAAMKANADETITALPEGEYFNWGORFPLISGGOKRIATARAAYVSPKILLIDEAT	604
Db	500	-IEANVAEANAYDEIMKLPHKFPDLVEBERQAOLSGGOKRIATARALVRPKILLIDEAT	558
QY	605	SALDTKSEGVVQAALERBAERGTITYIAHRLSTIKTANINIVLVNCKIAEOGTHDELVD	664
Db	559	SALDTSEAYVQVALDRAKRGRTITYIAHRLSTVRNADYIAGEPDGVYVERKGNHDELMKE	618
QY	665	GGATRKIYENORINEOKREDALEDAEDLTMNDIAIKTASASSPDLDCKPTTIDRTGT	724
Db	619	KGITFKLVTQTAGNEVE---LENA-----ADESSEIDALEMSSNDSSRLTRRST	668
QY	725	HKSVSAILSKRP-----PETPKYSLTILKLFVASFNRPEIPLYMLIGLVEVSLAGG	777
Db	669	RRAVRGSAQADRKISTEKDEALDESIPPSFMRIM----MLTEPRYVVGVEFCATINIGGL	724
QY	778	QPTQAVLYAKAISPLSLPESQYSLRHDAPFWSLIMFVVGIIQETIIOSTINGAFAVCSER	837
Db	725	QPARAIIISFKIIGVETRIDDPETK-RONSNLFLSLFALALIISEITPEFLOGGFEKGAEI	783
QY	838	LIRARASTAFETILIRODIAFPEKRENSGTALSTLSTETHLSGVSQVTGTLIMSTTL	897
Db	784	LTKRLRITWTFMSMRKODVSWPDDPKNTGALTATKRLADNAOVGAIGSRLAIVTONTANL	843
QY	898	GAAILIATLGMKIALVCIISVVPYLLACGFYRFMYLMFOFORSKLAYEGSANFACEATSS	957
Db	844	GTGIIISPYGMQTLTLLIAIVPILAIAGVEMKMLSGOALMKKKELEGAKIATAEIAEN	903
QY	958	IRTYAASLTREDWEIYHAQIDACGFRSLISLVRSLLIYASSQALVEFYCALGFWGCTL	101
Db	904	FRTYVSLTQOEKFEHMAQOSLOVYRNRSLRKAHJFGTTFSETOAMMFYSVAGCFREFGAYL	963
QY	1018	LGHEHYDIFREFVCESEILFEQASAGVFEFAPDMGAKANAAAEFRRLPFRKPOIDMWSE	107
Db	964	VAHKLMSIEDVLLVFSAVYVGGAAVGOVSSFAPDYAKAKISAAHIMIITEKPLIDYST	102
QY	1078	EGERLETVGEIEFRNRYFRPTREOPVLRGLDLTVKPGQYVALVPSGCGKSTYATL	113
Db	1024	EGLMPNTEGVRVGEVFNVPTRPDIPYLOGLSLEVKKQOTIALVSGSCGKSTVQLL	108
QY	1138	ERFIDALAGSLTVNGKROISKLININYSFSLVSOEPTLYOGTITKENIILGIVEDVPEE	119
Db	1084	ERFDPDLAGKVLIDGKRIKRLNQMWMLAHJIGVSOEPLIDPDCSTAEINIAAGDMSRVVSOE	114
QY	1198	FLIACDKDANIYDITMSLEPEFNTVVGSGGMLSGGOKRIATARAALVRPKILLIDEAT	125
Db	1144	EIVAAAEANIHATIESLPKYSITKYVDKQTSOLSGGOKRIATARALVRPHILLIDEAT	120
QY	1258	SALDSESEKVVQAALDAANGRTITYIAHRLSTIOKADIVYEPDQKIVESGTHSELVOK	131
Db	1204	SALDTESEKVVQOVALDRAKRGRTITYIAHRLSTIONADLIVYFONGRYVKEHGTHQOLLAQ	126
QY	1318	KGRYVELVNOS 1329	
Db	1264	KGITFMSVSVQA 1275	

RESULT 3
US-09-866-866A-2
; Sequence 2, Application US/09866866A

```

: Patent No. US20020102244A1
: GENERAL INFORMATION:
:
: APPLICANT: Sorrentino, Brian
: APPLICANT: Schuetz, John
: TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
: FILE REFERENCE: 1340-1-021CIP2
: CURRENT APPLICATION NUMBER: US/09/866,866A
: CURRENT FILING DATE: 2001-08-30
: PRIOR APPLICATION NUMBER: 09/584,586
: PRIOR FILING DATE: 2000-05-31
: PRIOR APPLICATION NUMBER: PCT/US99/11825
: PRIOR FILING DATE: 1999-05-27
: PRIOR APPLICATION NUMBER: 60/086,988
: PRIOR FILING DATE: 1998-05-28
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 1280
: TYPE: PRT
: ORGANISM: Homo sapiens
:
: US-09-866-866A-2
:
: Query Match          37.0%; Score 2489; DB 10; Length 1280;
: Best Local Similarity 42.1%; Pred. No. 5,9e-177;
: Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9

```

[illegible]

QY 778 QPQAVLYAKAISTLSPESQYSKLRHADFWSLMFFVVGIIOTITOSTGAAFPVCSER 837
 DB 725 QPFAAIFSSKIIIGVFRIDDPETK-RQNSMLFSLFLALGIIISITFELOGFFGKAGEI 783
 QY 838 LURASTAFRTLLRODIAFDKENSNGALSTFSLSTETKHLGSGVLTGLTMTSTL 897
 DB 784 LTRRLKRYMFRSLRODVSPFDPKNTGALTTRLANDAAQVGAIGSRLAVTQNIANL 843
 QY 898 GAAITIALIGMKLALVCISVPEVLLACGFYRYMLAQOFSRSLAVEGSANFACETS 957
 DB 844 GTEGIIISITFQMQLTLLALVPIIALIAGVEMKMLSGALKDKKELEGAKIATEAIEN 903
 QY 958 IRFVASTERDVMELIYHQAQRTSLISVLRSSILYASSQALVFCVALLGFWGGL 1017
 DB 904 FRTVASTQOEKHEHYAOSLOQVYRNSLRKAHIFGITSFTQAMMFYSYAGCFRGAYL 963
 QY 1018 LGHEEDIFREFVCFSEILFQAOSAGTVSFAPDMGKAKNAAEFRRLPQKQDQINWSE 1077
 DB 964 VAHKLMSFEDVLYFSAVYFGAMAVGVSSFPADYAKAKISAAHIIIMITKTLIDSYST 1023
 QY 1078 EGEKLEVEGELEFRVHRYTPRPOVYLRGLDLYPKPOYVALVPSGCGSTTALL 1137
 DB 1024 EGIAPMTLBNVTFGEVFNYPTRDIPVLOGLSTLEVKKGQTLALVSSGCGSTVQVL 1083
 QY 1138 ERYDVAISILVDGDISKLNINSTRSLVSOEPTLYOGTITENILIGIYEDVPEE 1197
 DB 1084 ERYDVLAKRVLIDGKEIRLNVQMLRAHGIYSOEPIILFDCSIANINAYGDNRSVYSC 1143
 QY 1198 FLTKACKDANIYDFMSLPEGEFTVYVSGKMGSGGOKORVATARALLRDKPILLIDEAT 1257
 DB 1144 EYBRAKENINAHFISLBRKXISTKYGDGKTQSLSGGOKORIALARLVRPILLIDEAT 1203
 QY 1258 SALDSEKRYVQAALDAAARGTTIAVAHRLSTIQADVIYVEDQKRYESGTHSELVOR 1317
 DB 1204 SALDSEKRYVQALDARBGRFCIVIAHRLSTIQADVIYVEDQKRYESGTHSELVOR 1263
 QY 1318 KGRYVELVMS 1329
 DB 1264 KGRYVELVMS 1275
 RESULT 4
 US-09-769-097-4
 ; Sequence 4, Application US/09769097
 ; Patent No. US2002005128A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kimberly Anne Brun
 ; APPLICANT: Richard James Chenery
 ; APPLICANT: Harna Ellens
 ; APPLICANT: John Anthony Feld
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
 ; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
 ; FILE REFERENCE: GP-50009-C2
 ; CURRENT APPLICATION NUMBER: US/09/769,097
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 09/208,809
 ; PRIOR FILING DATE: 1998-12-09
 ; PRIOR APPLICATION NUMBER: 09/156,800
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: US99/20770
 ; PRIOR FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 1272
 ; TYPE: PRT
 ; ORGANISM: RATTUS RATTUS
 US-09-769-097-4
 Query Match 37.0%; Score 2487.5; DB 10; Length 1272;
 Best Local Similarity 42.0%; Pred. No. 7.6e-177;

Matches 534; Conservative 226; Mismatches 461; Indels 49; Gaps 10;
 QY 90 KQLEIEIKVNISEFGMLRYATKMDILIMVISTCAIAAATFORIM----- 136
 DB 23 KKEKKKPAVSVLTMERVAGWLDREFYMLIGTLAIIHIGLALMLVFECDMDSPANNG 82
 QY 137 -----YQIS--YDEYFDELTKNVLKFNLYVLGIGEFYTVVSVGFIYGEHATOKIREY 187
 DB 83 NNRSMSTFYNATDIYAKLEDMTYAAVYTGIGAGVLIIVAIQVSTWCLAARQIHRIRQK 142
 QY 188 YLESIIRONIGYDKDGAGVETTRTADNTLIDGISEKGLTLLATAPVFTIAYK 247
 DB 143 FFAIIMNOEIGMDVADVGEMLNTRLDVDSKINEGIDKIGMFQMAATFGGFIIGFTR 202
 QY 248 YMKALICSTTIALVLTMGSGSOFIIRKSKSIDSYGAGTVAEEVSSIRNATAFGR 307
 DB 203 GMLTLVILIASPVLGSLGIMAKLISFDFKLOQVAAKGNAAVEVLAIRVIAFGG 262
 QY 308 DKLAKQEVHLDAEKWGTKNQIVMGPMIGAMGLMYSNGIGFMGSRFLVDGAVDGD 367
 DB 263 KKELEERYNNMLEEAKRLGIKAITANISMGAAFLLIYASVALAFYGTSLVSKETYG 322
 QY 368 ILFVLMALIGSFLGSLGNVSPNAQAFINAAAKIGITIDROSPLDIPYSEGTLDHFE 427
 DB 323 VLTVEFSLVIGAPSVQASPNIEAFANARCAAEVFSIIDNKRSIDFSKSGHKNPNI 382
 QY 428 HIELRWVKHIYPRPEVTVMEDVLSMPAGKTATVAGSPSGSTVAGYERFVMPVGR 487
 DB 383 NLEFKNIHFSYPRKDVOLIKGLNLKLVKSGQVATVAGNSCGGSTVOLLQRLYDIEGE 442
 QY 488 VLLGDHIDKDLNRLRQOISLVSOEPLVFGTTIYNIRHGLIGTYENESDKVRELIE 547
 DB 443 VSIDGODIRTNVRYLREIIGVSOEPLVFATIAENIRYG-----RENTYME----- 493
 QY 548 NAKMANADEFTALPEGETVNGORGFLLSGGOKORIVARVSDPTILLIDEATSL 607
 DB 494 KAYKEANAYDFIMKLPHKEDTLVGERGAQSLSGGOKORIALARLVRPILLIDEATSL 553
 QY 608 DTKSEGVQAALERAEGRTTIVIAHRLSTIKTANHIMVYLVNKGIAEGTHDELVDGGA 667
 DB 554 DTSEFAVQAALDKAREGRTTIVIAHRLSTVNRADVIAAGPDGVYIEQGNHDELMKEGI 613
 QY 668 YRKLVEQQRINEKEA--DALEDADADLTNADIAKTAASSASDLPDPTTIDRTGTH 725
 DB 614 YERLVMTQAGNIEELGNCESSKIDID--NVDMS--SKDSSSLIRRRSTRKSIRGPH 668
 QY 726 KVSASALISKRPEPTPKYSLMTLTKFVASFNPEIPYMLIGLVFVSGGQPOAVLY 785
 DB 669 DQDGEIETKEALDDVPPASFWRIK-----LNSTEMPYVGVFCALINGGLDPAFSTIF 724
 QY 786 AKAISTLS--LPEQYSKLRHADFWSLMFFVVGIIOTITOSTGAAFPVCSERLIRRA 842
 DB 725 SKYGVFTKNDPBEIQ-----RQNSMLFSLFLALGIIISITFELOGFFGKAGEILTREL 780
 QY 843 RSTAFRTIIRODIAFDKENSNGALSTFSLSTETKHLGSGVLTGLTMTSTGAAIT 902
 DB 781 RYMWFSKMLRODISWDPDKNTGALTTRLANDAAQVGAIGTSRLAVTQNIANLSTGII 840
 QY 903 IALAIKMLALVCISVPEVLLACGFYRYMLAQOFSRSLAVEGSANFACETSIRTYA 962
 DB 841 ISLIYQMQLTLLALVPIIALIAGVEMKMLSGALKDKKELEGAKIATEAIENRIVY 900
 QY 963 SLTREDVWEIYHQAQRTSLISVLRSSILYASSQALVFCVALLGFWGGLTLLGHE 1022
 DB 901 SLTREQKFTMYAQSLOIYRNAKKAHFGITFSTQAMMFYSYACRPAIYVAREL 960
 QY 1023 YDIFRFVCFSEILFQAOSAGTVSFAPDMGKAKNAAEFRRLPQKQDQINWSEGERL 1082
 DB 961 MTFENVLIVFSAIVFGAMAVGVSSFPADYAKAKISAAHIIIMITKTLIDSYSTGLKP 1020
 QY 1083 ERYGEIEFRNRYTPRPOVYLRGLDLYPKPOYVALVPSGCGSTTALLERFD 1142
 DB 1021 NMLEGVAKENGVMFNYPTRDIPVLOGLSTLEVKKGQTLALVSSGCGSTVOLLERFD 1080


```

: PRIOR APPLICATION NUMBER: 09/584,586
: PRIOR FILING DATE: 2000-05-31
: PRIOR APPLICATION NUMBER: PCT/US99/11825
: PRIOR FILING DATE: 1999-05-27
: PRIOR APPLICATION NUMBER: 60/086,988
: PRIOR FILING DATE: 1998-05-28
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: Patent version 3.0
: SEQ ID NO 4
: LENGTH: 1280
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-866-866A-4

Query Match      36.9%  Score 2480;  DB 10;  Length 1280;
Best Local Similarity 42.0%;  Pred. No. 2.8e-176;
Matches 534;  Conservative 229;  Mismatches 449;  Indels 60;  Gaps 9;

QY 94 EEIKVNISFPGIMRYATKMDILIMVISTICALAASTFORIMLY----- 137
DB 28 KKKKPTVSFVSFMRYSNMILDKLYMVGTLAAIHGGLPLMLVGEEMTDIFANAGNLED 87
QY 138 -----QISYDEFY-----DELTKNVLVYVYLGIGEFVYVYSTVGFITYGEHATOKI 184
DB 88 LMSNTNRSDINDTGFPMLEEDMTRYATYSGIGAGVVAAYIQVSPCLAGROIHKI 147
QY 185 REYVESLIRONIGYFDKIGAGEVTRITADTNLIDGISEKVGILTLTALATFVAFIIA 244
DB 148 RKOFHAIKROEIGFWDVDELNTRLTDVSKIWEVIGDKIMGFQSMATFFETGFIYG 207
QY 245 YKVKMLALICSTVALVLTMGSGGSQIIRKSKSLDSYGAGGYAEVVISIRATAF 304
DB 208 FTRGKMLTLVLAISPVLGLSAAVAAKILISFTDELLAARAGANAEEVLAIRIVIAF 267
QY 305 GTODKLAEQYEHVLEDEAEKGTKNQIYMGFMIGAMEGLMYNGJGFMMGSRFLVDGAVD 364
DB 268 GGOKKELEBRYNKNLEAKRIGIKKAITANISIGAFELITYASTALAFWGTILVLSGETS 327
QY 365 VGDILTVMALIIISFSLGANSVPMNAQFTNAVAAAACIFGTIDROSPLDPYSNEGRTIDH 424
DB 328 IGOVLTFFSVLIGAFSPVQASPSIEAFNARGAAVEIKIIDNKRPSIDYSKSGHKPN 387
QY 425 FEHGLELNKHIYPSRPVYMEVSVLSMPAGKTTALVGPESGSKSYVGLVERMYMV 484
DB 368 IKGNELEFRNVEHSYSPSRKVKILKGLNKLKVGSGQVVALVNGSGCKSTVOQMORTYDPT 447
QY 485 KGVLLDHDKIDNLRMLRQOISLVSEPPVYFGTTIKYKINHGILGTRENESEDKYAE 544
DB 448 EGMVSVQDQDRTIIVRFLREIIGVSOEPVLFATTIAENIRYQ-----RENVTYDE--- 499
QY 545 LIENAKANAAHDFITALPEGETYVNGORGLLSGQKORAIARAIVSDPKILLDEAT 604
DB 500 -IEKAVKRNAAVDFIMKLPKHFDTLVGERGALSGQKORAIARAIVNPKRILLDEAT 558
QY 605 SALDRKSGVGOALERAAGRTTIVIAHRLSTIKTANHIVLVNGKIAEGSTHEDVDR 664
DB 559 SALDSESEAVVOALDKRKRGTTIVIAHRLSTIKTANHIVLVNGKIAEGSTHEDVDR 618
QY 665 GAVYKRLVEAQRINEOKADALEDAEDLTNAADIAKIKTASSASSDDLGKFTTIDRTGT 724
DB 619 KGIYFKLVTMTAGNEVE---LENA-----ADESKSEIDALEMSNDRSRLRKST 668
QY 725 HKSVSSAILSKRP-----PETTPKYSILWTLKLFVASFNRPEIPYMLIGLVFSLVAGG 777
DB 669 RRSYVSGOQORKLSTKEALDESIPVSFWRIMK---LNLTEPMYFVAVGCALINGL 724
QY 778 QPTQAVLAKAISTLSLESOYSKIRHADQWSLMEFVVGIIQFTTOSTNGAARVAVCSER 837
DB 725 QPAVALISKTIIGVTRIDDETK-RONSNNFSLFALGIIISFTTFLOGTFKAGEI 783
QY 838 LIRARASTAFITILRODIAFDPKRENSGALTSFLSTEFKHLSGVGYTIGILMTSTTL 897
DB 784 LTKRLRYVFSMLRKQDVSWEDDPKNTITGALTTRILANDAAOVKGAIGSLAVITONIANL 843

```

```

QY 898 GAIIITAIAGMKLALVCISVPEVLLACGFYRPMYLAOFOSRKLAYEGSAPACETATSS 957
DB 844 GGIILISIFYGQMLTLLLAIVPIIALAGVEKMKLSGALKDKKLEAGKATATAIEN 903
QY 958 IRTVASLFRERDVWEIYHQAOLDQGFSTLSIVLRSSLIYASSQALVFCVALCFWYGTLL 1017
DB 904 FRTVASLTOBOKFEHMYAOSLQVPRNLSLRKAHIFGIFTSFTQAMMYFSYAGCFRGAVL 963
QY 1018 IGHHEVDIFRFVCESEILIFGAQSAGTVSFAPDQKAKANAAEFRLDRKPDINMSE 1077
DB 964 VAKKLMSPEVDLVFSAVAVGAAVQVSSFADYAKAKATISAAHITMIIEKPLDLSYST 1023
QY 1078 EGEKLETVGEIEFRNVHVRPTPEOPVLRGLDVLVKGQYALVPGCGGKSTTIAL 1137
DB 1024 EGMPTLELGNVTFGVYVNTPRPIPLQGLSLEVKKGQTLALVSSGCGKSTYVOLL 1083
QY 1138 ERFPYDAGSILVDGDISKLNINSTRSFLSLVSOEPFLTYOGTIKENILGIVEDVPE 1197
DB 1084 ERFPYDLACKVLLDGKEIRLNVQWLRHAGIYVSOEPILFDCSIABNIAYGDNRSVSGE 1143
QY 1198 FLIKACKDANIYDEIMSLPEGNTVYSGKMGISGQKORVAAARLLDPRKILLDEAT 1257
DB 1144 EYRAKKEINHAIFESLSPKSTFKVGDGTOLSGQKORAIARALVQPHILLDEAT 1203
QY 1258 SALDSESEKVVQALDAAARGRTTIAVHRLSTIQKADVIYFEDQKIVESGTHSELVOK 1317
DB 1204 SALDSESEKVVQALDAAARGRTTIAVHRLSTIQKADVIYFEDQKIVESGTHSELVOK 1263
QY 1318 KGRYFELVWQS 1329
DB 1264 KGIYFSWVSQA 1275

```

RESULT 7

```

US-09-769-097-2
: Sequence 2, Application US/09769097
: Patent No. US20020055128A1
: GENERAL INFORMATION:
: APPLICANT: Kimberly Anne Brun
: APPLICANT: Richard James Chenery
: APPLICANT: Harna Eilens
: APPLICANT: John Anthony Feild
: APPLICANT: Lin Yue
: TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
: FILE REFERENCE: GP-50009-C2
: CURRENT FILING DATE: US/09/769,097
: PRIOR APPLICATION NUMBER: 2001-01-25
: PRIOR APPLICATION NUMBER: 09/208,809
: PRIOR FILING DATE: 1998-12-09
: PRIOR APPLICATION NUMBER: 09/156,800
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: US99/20770
: PRIOR FILING DATE: 1999-09-10
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 1272
: TYPE: prt
: ORGANISM: RATTUS RATTUS
US-09-769-097-2

```

```

Query Match      36.8%  Score 2475.5;  DB 10;  Length 1272;
Best Local Similarity 42.0%;  Pred. No. 5.9e-176;
Matches 533;  Conservative 225;  Mismatches 463;  Indels 49;  Gaps 10;

```

```

QY 90 KTOLEIKVNISFPGIMRYATKMDILIMVISTICALAASTFORIMLY----- 136
DB 23 KKKKPTVSFVSFMRYSNMILDKLYMVGTLAAIHGGLPLMLVGEEMTDIFANAGNLED 82
QY 137 -----QISYDEFYDELTKNVLVYVYLGIGEFVYVYSTVGFITYGEHATOKIREY 187

```

Db 83 NNRSMFYNAADTYAKLEDEBMAITYAYTNGCAVLIVAYIOVSLMCLAAGROIHKIRK 142
 QY 188 YLESLIRONIGYEDKLGAGEVTRITADNLIDGISEKGLTLTALAFVTAIYAVR 247
 Db 143 FFAIINNOIGEMDVHDVDELNRLDDVSKIINGIDKIGMFQOMATFFGFIIGTR 202
 QY 248 YMKALICSTYVALVLMGGGSOFTIKYSKSLDSTGAGGVAEEVYSSIRNATAFGTQ 307
 Db 203 GMLTLYIILAIISFVLGSLAGIMAKILSSFTDEKLOAVAKAGAAVEVLAIRVIAFGCQ 262
 QY 308 DKIAKGEVHIDAEKMGTKNOIVMGFMIGAMGLMYNSNGLGFMWGSRLVYGAVDVC 367
 Db 263 KKELERYNNNLEAKRGLIKKALTANISMGAAFLLIYASTALAFWGTSLVISEKTYITGQ 322
 QY 368 ILTVLAILIGSFSLGNVSPNAQFTNAAVAAKIFGTIDROSLDPLDPSNEKTLDFHEG 427
 Db 323 VLVFVSVLIGAFSVQASPNIEAFANAGAAEVEFSIIDNKSIDSFSKGRPNIOG 382
 QY 428 HIELRNKHIYSPREVTYVEDVSLSPAGKTTALVPGSGSKSTVVGLEVFYMPVGT 487
 Db 383 NIEFKNIHFSYPSRKDVQILKGLNLKYSGQVALVNSGCGSKSTYQLQRLYDPIEGE 442
 QY 488 VLLDGDIDNLNRLRQOISLYSOEVLFGTTIYKNIHRLGLGTRENSDEKVELIE 547
 Db 443 VSDIGODIRITINRYLREITIGVSOEVLFAITIAENIRG-----RENYTME---IE 493
 QY 548 NAKMANADFTALPEGETYNGVORGFLLSGQOKORAIARAIVSDPKILLDEATSA 607
 Db 494 KAVKEMAYDFIMKLPKHKDTLVGERGAQLSGQOKORAIARAIVNRPKILLDEATSA 553
 QY 608 DPKSEGVQALERAEGRTTIVIAHRLSTIKTAHNIIVLVNKGIAEQTHDELVDRG 667
 Db 554 DRESEAVVOAALDKAEGRTTIVIAHRLSTVRNADYAGDGVIEQGNHDELMREKGI 613
 QY 668 YRRLVDAQHINEQKA--DALEDADADLTNADIAKIKTSSASDLDGKPTIIDRGTH 725
 Db 614 YELVMTQAGNIEILGNECESKSDID--NVMS--SKDSQSLIRRSTKIRGPH 668
 QY 726 KVSASALSKRPETPKYSIMTLKVFVAFNPEIPMYLIGLVFVLAGGQPTQAVL 785
 Db 669 DDDGELSTKALDDVPPASFWIRLK-----LNSTEMPYFVVGFCALINGLOPABSIT 724
 QY 786 AKAISTLS---LPESQSKLRHDADFWSLMFVVGIIQFTVSTNGAFAVCEBRLIRA 842
 Db 725 SKVAVGFTKNDPEIQ-----RONSNLFSILFLIIGIISFTFLOGTFEGKAGBIILKRL 780
 QY 843 RSTAEFTIIRODIAFEDKENSFGALTSPLSTFKHLSVSGVTGLGIIIMTSTTGAAIT 902
 Db 781 RYVWFKSMRQDISWDDPKNTTGALTTRLANDAAOVKGAWSRLAVITTONIANIAGTII 840
 QY 903 IALAIKMLALVCISVVPVLLACGFYFYMLAOFOSRSKLAYESGANFACETASIRTA 962
 Db 841 ISLIYMOQLTLTLATVPIAINGVEMKRLSGOALKDKELESGKIALEALENFTVY 900
 QY 963 SLTREBDVWEIYHAQIDAQGRSLISVNSLLYASSQALVFCAVLFVGGTLLGHE 1022
 Db 901 SLTREKFEEMZYQSIQIPIYRNALKRAHYGIFESFQAMMYFSYAACFRDVIYAREL 960
 QY 1023 YDIFRFVCESELIFGAQSGVTFSPAPDKGAKNAAEERLDFDRKPOIDNNSEBEKL 1082
 Db 961 MPEENLVFSAIVFAMAAGOVSPAPDYAKAKVASHIRIETEKIPEDISSTEBELK 1020
 QY 1083 EYVEGEIEFFANVHRRPTREPOVLRGLDITVKGQVALVPGSGSKSTYIALLERFYD 1142
 Db 1021 NMLEGVAKFNGVAFNPTRNIPVLOGLSLEVAKKGOTIALVSSGCGSKSTYVOLLERFYD 1080
 QY 1143 AINGSLVLDGKDISKLNINSYRSFSLVSOEPLLVOGTIKENILLGIVEDVPEELIKA 1202
 Db 1081 PMAGTFFLDGKEIKOLNVOALRAHLGIVSOEPLIFCSISENIAYAGDNSHYVSHKEIVKA 1140
 QY 1203 CRANVYDFIWSLPBGFNTYVSGKGLSGGQOKORAIARALLRDPKILLDEATSA 1262
 Db 1141 AKBANIHOFIDSLPERKYNTRVSGKGQOLSGQOKORAIARALVROPILLDEATSA 1200

QY 1263 ESEKVOAALDAAARGTTIAVAHRLSTIQADVIYVFDGAKIVESGTHSELVOKKGRY 1322
 Db 1201 ESEKVOAALDAAARGTTIAVAHRLSTIQADVIYVFDGAKIVESGTHSELVOKKGRY 1260
 QY 1323 ELVNLQSLGK 1332
 Db 1261 SMVSVQAGAK 1270
 RESULT 8
 US-09-866-866A-8
 ; Sequence 8, Application US/09866866A
 ; Patent No. US20020102244A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sorrentino, Brian
 ; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
 ; FILE REFERENCE: 1340-1-021CIP2
 ; CURRENT APPLICATION NUMBER: US/09/866,866A
 ; CURRENT FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: 09/584,586
 ; PRIOR FILING DATE: 2000-05-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/11825
 ; PRIOR FILING DATE: 1999-05-27
 ; PRIOR APPLICATION NUMBER: 60/086,988
 ; PRIOR FILING DATE: 1998-05-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 1276
 ; TYPE: PRF
 ; ORGANISM: Mus musculus
 US-09-866-866A-8
 Query Match 36.8%; Score 2475.5; DB 10; Length 1276;
 Best Local Similarity 41.4%; Pred. No. 6e-176;
 Matches 527; Conservative 230; Mismatches 466; Indels 49; Gaps 6;
 QY 90 KTOLEELKIVNISFGMLRATKMDILMIVSTICATAAASIFQIMX----- 137
 Db 23 KKEKERRKPAVSVLTMRVAGWMDRLYMLVGTALAIHGAALPMLIFGDMTDSFASVG 82
 QY 138 -----QISYDFEYDELTKNVLVFNVLGICEFVTVSVTGFTYGEHATOK 183
 Db 83 NVKSNSTNMSEADKRAFALEEBMTTYAYYIGIGAGVLIYVAYIOVSFMCLAGROIHK 142
 QY 184 IREYIESTLRONIGYEDKLGAGEVTRITADNLIDGISEKGLTLTALAFVTAIY 243
 Db 143 IROKFFHAIINNOIGEMDVHDVDELNRLDDVSKIINGIDKIGMFQOMATFFGFI 202
 QY 244 AYVYKMKALICSTYVALVLMGGGSOFTIKYSKSLDSTGAGGVAEEVYSSIRNATA 303
 Db 203 GFTRGWLTLYIILAIISFVLGSLAGIMAKILSSFTDEKLOAVAKAGAAVEVLAIRVIA 262
 QY 304 FGTODKIAKGEVHIDAEKMGTKNOIVMGFMIGAMGLMYNSNGLGFMWGSRLVYGAV 363
 Db 263 FCGGKKELERYNNNLEAKRGLIKKALTANISMGAAFLLIYASTALAFWGTSLVISEK 322
 QY 424 HFEHIELRNKHIYSPREVTYVEDVSLSPAGKTTALVPGSGSKSTVVGLEVFYMP 483
 Db 383 NIEGKNIHFSYPSRKDVQILKGLNLKYSGQVALVNSGCGSKSTYQLQRLYDPI 442
 QY 484 VRLVLDGDIDNLNRLRQOISLYSOEVLFGTTIYKNIHRLGLGTRENSDEKVEL 543
 Db 443 LDGKVSIDGODIRITINRYLREITIGVSOEVLFAITIAENIRG-----RENYT 493
 QY 544 ELIENAKMANADFTALPEGETYNGVORGFLLSGQOKORAIARAIVSDPKILLDEA 603

```

Db 494 DEIEKAVKAAVDFIMKLPDHOEDLVGRGAHVSQGOKORAIARALVNRKILLDDA 553
Qy 604 TSALDTKSEGVVQAALEERAEGRTTIVIAHRLSTIKTANHIVLVNGKIAEGSTHDELYD 663
Db 554 TSALDTSEBAVVOALDKAREGRTTIVIAHRLSTVNRADVIAGFDGVIVEGDNHDELMR 613
Qy 664 RCGATKRLVEARINQOKADALEADADLNNADIATKKTASSASSDDGKPTTIDRIG 723
Db 614 EKGITFKLMTQTAGNEILGNBCKSKDEIDNLDMS---SKDSGSSLLRRSTRKSIG 670
Qy 724 THRSVSAIISKRPEPTPKYSILMTLLKFVASFNRPEIPLYMIGLVSVIAGGOTPAY 783
Db 671 PHDQBRKSTKRALDEVDVPASFWRILK----LNTSEMPYFVVGICAILINGLOAPESY 726
Qy 784 IYAKAISTLS---LPESQYSLKRDADFWSLMFVVGIIQFTQSTNGAFAVCSERLIR 840
Db 727 IFSKVGVFTNGGPETQ-----RONSMLFSLFLILGILSIFTEFLQGFPGKAGEILRK 782
Qy 841 RARSTAFRTILRODIAFPCKEENSGALTSPLSTERKHLISVSGVTLGITIMSTTLGA 900
Db 783 RLRVAFKSMLRQDVSWEDDPKNTGALTTRLANDAAQVKGATGSHLAVIFONIANILGTG 842
Qy 901 IIALAIGKRLALVCISVVPVLACGFYRYMLAQFOSRSKLAYEGSANFACEATSIRT 960
Db 843 IISILYGMQLLILALAIYPIAIAGVEMKMLSGOALKDKELBESGKIAIETAIENFRT 902
Qy 961 VASLTERDWEIYHAQDLQAGRTSLISVLRSSLLYASSQALVFECVALGFWYGGTLLGH 1020
Db 903 VYSLTREQEFTMYAOSLIQIPYRNMRKKAHVGFITFSFOAMMYFSYACFRGAVLYNQ 962
Qy 1021 HEYDIFRPFVCGSEILFGQASAGIYFSPAPDMGKAKNAAEERRRLFDKRPQIDNMSSEE 1080
Db 963 QLMTEENVLVSAIVFGMAVGOVSEFAPDYAKATVASHIRIETKEPDISYSTOGL 1022
Qy 1081 KLEIYEGEIEFRNHFERYTREPQVLRGLDLTVRPGQVALVGPSCGKSTTIALIERE 1140
Db 1023 KPNMLEGNVQFSGVFNFTYRPSIYLOGLSLEVKKGQILALVSGSCCKSTVOLLIERE 1082
Qy 1141 YDAIAGSLVDGKDISKLNINSYRSLVSQEPFLYOGTIFENILGLIYEDVPEEPLI 1200
Db 1083 YDPMAGSVFLDKEIKQLQVQMLRAOLGIVSQEPILFDCSIENATVAGDNSRVSYEEIY 1142
Qy 1201 KACKRANLYDFIMSLPEGNRYVVSQKGMLSGQKORVAAIALRDPKILLIDETSAL 1260
Db 1143 RAAKRAANHQFIDSLPDKNTYRGDKGTQLSGQORALAAIALVRQPHILLIDETSAL 1202
Qy 1261 DSESEKVVQAAADAAARGTITVAHRLSTIQADVIYFVDGKIYESGTHSELYOKGR 1320
Db 1203 DTESKVVQEAALDKAREGRTTIVIAHRLSTIONADLIYVIONGKYKEHGTHOOLAOKCI 1262
Qy 1321 YELVNLQSLGK 1332
Db 1263 YFSMVSVOAGAK 1274

```

```

RESULT 9
US-09-866-866A-6
Sequence 6, Application US/09866866A
Patent No. US20020102244A1
GENERAL INFORMATION:
APPLICANT: Sorrentino, Brian
TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
FILE REFERENCE: 1340-1-021CIP2
CURRENT APPLICATION NUMBER: US/09/866, 866A
CURRENT FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: 09/584, 586
PRIOR FILING DATE: 2000-05-31
PRIOR APPLICATION NUMBER: PCT/US99/11825
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: 60/086, 988
PRIOR FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 27

```

```

SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 1276
TYPE: PR
ORGANISM: Mus musculus
US-09-866-866A-6

Query Match      36.5%; Score 2453.5; DB 10; Length 1276;
Best Local Similarity 40.9%; Pred. No. 2,6e-174;
Matches 520; Conservative 232; Mismatches 461; Indels 57; Gaps 7;

Qy 90 KQLEIKVNIISFPGMRATKMDILIVYISTICAIAASTFORIML----- 136
Db 23 KKEKKKKPAGVGFGRADWLKCMILGTLAIIHCTLLPLMLVGNMTDSEPTKAE 82
Qy 137 -----YQSYDEFYDELKKNLYEFYLLGIEGFVYVYVSGVITYGERA 180
Db 83 ASLISITNOSGPNSTLIISNSLSLEEMAIYAYYIGICAGVLIYVIOVSLMCAAGKQ 142
Qy 181 TOKIREYLESILRONIGYFDKLAGEVTRITADTNLIQDISEKVGTLTALATFVA 240
Db 143 IHRKQKFFHAIINQEIIGFVDHVDGELNTRLDVSKINDIGDKIGMFQSITFFLAG 202
Qy 241 FIITAYKYKRLALICSTVALVLTIMGSSOFLIKSKSLDSYGAGTVAEVISSIRN 300
Db 203 FLIGFISGKRLVLVLAVPLIGLSALMAKVLTSPTNKELOAYAKAGAVAEVLAIR 262
Qy 301 ATAFGTQDLAKOYEYHDEAEKQKNOIYMGFMIGAMEGLMYSGIFMMGSRFLVD 360
Db 263 VIAFGQOKELERYKNLNEAKNVGIKKAITASISIGIYLVLYASIALAFMYGSLVIS 322
Qy 361 GAVVDGILITVLMALIGSFSLGANSPPNAQAFNAAVAAAKIFGTIDROSPDPPSNEGK 420
Db 323 NEYSIGEVLTVFSSILGLGFSIGHLAPNIEAFANRGAFAELFKIINDPSIDSTKCY 382
Qy 421 TLDHEGHELANVKHITYRREYVYMEDVYSMPAGKTALVGPSCGKSTVGLVERE 480
Db 383 KPDISLNGNLEFNHFNHFNYSRSEVOILKGLNLKYNVSGQTVALVNGSCGKSTVOIMQRL 442
Qy 481 YMPVRGTVDLHDHIDKIDLMRLQOISLVSQEPVFGTITKKNRHGILGKYNESHD 540
Db 443 YDPLEGVSIDQDITITVIRKILREILIGVSOBPVLFATITENITNG-----RED 493
Qy 541 KYRELIENAKKANAHDFITALPEGETVNVQGEFLISGQORAIARAVVSDPKILL 600
Db 494 YTMDEIEKAVKAAVDFIMKLPDHOEDLVGERGAQLSGQOKORAIARALVNRKILL 553
Qy 601 DEATSLDTPKSGVVOALERAEGRTTIVIAHRLSTIKTANHIVLVNGKIAEGSTHDE 660
Db 554 DEATSLDTPSEBAVVOALDKAREGRTTIVIAHRLSTVNRADVIAGFDGVIVEGDNHDE 613
Qy 661 LVDRGAVKRLVEAQ-RINEOKRAD---ALDEDADEDLTNADIKIKTASSASSDLOK 715
Db 614 LMRKGIYFKLVMTQTRNGEIEPGNNAVYSGSDTDASELTSSEKSPILRIRSYVNRK 673
Qy 716 PTTIDRTGTHKSVSAIISKRPEPTPKYSILMTLLKFVASFNRPEIPLYMIGLVSVIAG 775
Db 674 QDOERRLSKAEAYD-----EDVPLVSFWRL-----NLMSPLYLVGLVLAING 720
Qy 776 GGQPOAVIYAKAISTLSLPESQYSLKRDADFWSLMFVVGIIQFTQSTNGAFAVCS 835
Db 721 CLOPFAIVFISRIYGVFSDDDHETK-RONCNLJLSFLFLVLMGLISVYVFFGQFFGKAG 779
Qy 836 ERLIRARSTARFTILRODIAFFDEENSGALTSPLSTERKHLISVSGVTLGITIMST 895
Db 780 ELITKRVIRMYKSMLRQDISFDDHKNSTGSLTTRLSNDASSVGAMCARLAATVONYA 839
Qy 896 TLGAAIIALAIGKRLALVCISVVPVLACGFYRYMLAQFOSRSKLAYEGSANFACEAT 955
Db 840 NLTGVIISLVYGMQUTLITLIVITPLIYAGIIEKKLSGQALKKQKOLEISGKILATEGI 899
Qy 956 SSIRTVASLTERDWEIYHAQDLQAGRTSLISVLRSSLLYASSQALVFECVALGFWYGG 1015

```



```

Db 293 EEMERARAREANAYDFIMEFPNKNFTLVGEKQAGMSGQKORIALAARLVANPRLILDEA 352
Qy 604 TSLADPKRSEGVQALERAEGRTTIVIAHRLSTIKTAHNIIVLVNCKIAEGOTHELDVD 663
Db 353 TSLADSESKSAVOAALEKASKGRTTIVIAHRLSTIRADILVLKQIALEKHAHEALMA 412
Qy 664 RCGAYRKRLVEAQRINQEKADALEDAEDLTADIAKIKITASSASDDLCKPTTIDRTG 723
Db 413 KRGILYSLVMSQDIKKADEDMESMTYSTERTKS--LPLHSVSKISDF-----IDK-- 462
Qy 724 THKVSASAILSKRPETTPKYSILMTLKEFVASFNRPEIPMLIGLVSVLGGAGGQPTQAV 783
Db 463 AESTQSKEIS-----LPEVSLKTLK---LNKPEMPVVLGTSLASVANGVHPVFSI 512
Qy 784 LVAKAISTSLPESQYSKLRHDAEFWSLMFEVVGIIQFTOSTNGAFAVCSRLIRRAR 843
Db 513 IFKAITMFG--NNDKTTTKHDAEISMIFFVLIGVICFVSFYMQGLFYGAGELTMRLR 570
Qy 844 STAFRTILRODIAFEDKENSGLTSLSTEDKHLGSGVTLGLTILMTSTLGAIIIT 903
Db 571 HIAFKMLVODIAMFDEKEKSTGTLTILADIAQIGATGSRIGVLTQONATNMGLSVII 630
Qy 904 ALAIGKMLAVCISVVPVLLACGFYFMYLQFOSRSKLAVESANFACETSISITVAS 963
Db 631 SFYIGEMETFLISTIAVPLAVTGMETAAWTFGANKOKELKAKIATFALENIITIVS 690
Qy 964 LRERBWEIYHAQLDAQGTSLISVLRSSLYASSOALVFCVALGFYWGTLGHHEX 1023
Db 691 LRERKAFEBMYEEMLOTOHRNTSKAKQIIGSCYAFSHAFYFAVAGFRGAVLIQAGRM 750
Qy 1024 DIFRFVFCSEILFGASQATVFSFAPDGMKAKANAAEFRLDRKPOIDNMSEBEKLE 1083
Db 751 TEBGMETVTAALVGMALIGKTLVLAPEYSKAKSGAAHLFALKEKKNIDSRQEBGKPPD 810
Qy 1084 TWEGEIEFRNHRYPTRBQPYLGLDLTVKBPQYVALVGPBGCKSTTIALLEFYDA 1143
Db 811 TEGNLEFEVSEFFYPCRPDVFILRLSLIERKTVAFVSGGCKSTSVQQLQLYDP 870
Qy 1144 IAGSLVNDKDISKLINSTRSFLSVSQEPITLYOSTIKENILGLVEDDVPPEFLIKAC 1203
Db 871 VOGQVLPEDVDAKELVQMLRSQIALVPEPVLFCNSIAENIAYGNSRVPVDEIKKA 930
Qy 1204 KQANITDFIMSLPEGEFTVYSGKMLSGQKORIVALAARLBDPKILLDEATSLDSE 1263
Db 931 NMANIHSEFLEGPEKNTQVGLKQALSGQKORIALAARLQKPKILLDEATSLDND 990
Qy 1264 SERVOALDAARGRTTIAVAHRISTIQADVIYVFDGKIVESGTHSELVOKGRYYE 1323
Db 991 SERVOALDKAKRTGRTCLVTHRLSAIONADLIIVLHNGKIKEGTHQHELLANRDYFR 1050
Qy 1324 LVNIQSL 1330
Db 1051 LVNAQSV 1057

```

RESULT 14

```

US-09-882-694-11
; Sequence 11, Application US/09882694
; Publication No. US20030009782A1
; GENERAL INFORMATION:
; APPLICANT: Davick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Crasta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; TITLE OF INVENTION: Detoxification
; FILE REFERENCE: 35718/208255
; CURRENT APPLICATION NUMBER: US/09/882,694
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1263
; TYPE: PRF
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 157
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-882-694-11

Query Match      28.9%; Score 1940; DB 9; Length 1263;
Best Local Similarity 36.8%; Pred. No. 5,6e-136;
Matches 490; Conservative 231; Mismatches 483; Indels 128; Gaps 30;

Qy 54 ADKHHPRSSNNVAVNEVDALIAHLPEDEQVLTQLEELKIVNS--PFGIMRATKM 112
Db 2 ADSEKPRRQDSSESS-----SHPPE-----KETEGSIDYLRIFRYADKY 44
Qy 113 DILIMVISTICAIAASTQRIIMLYQISY-----DEEYDELTKNLYEVYVL 158
Db 45 DMTLVNIALICALGASALPLMSIIFGSPFNKNYNSGDSPEAKADVDHFLMFVYL 104
Qy 159 GIGEFVTVVSYGFIYTGHAQKIREYILESLQNGICFPLKLAGEVTTITLDTNL 218
Db 105 FIKGFVLTIVSTAITITSAIRTRTLRNVFLTECLQDEWHFQKQNGAIAATYVTNGNR 164
Qy 219 IODGISEKVLITATATEVTAFTIAIVKWKALICSSFTYVALVTMGGSQFIITKYSK 278
Db 165 IOTGIANEKLVFTVQALSMFSAFVVALASQKALITMTVTPAIFLVTGICIAIDAQEA 224
Qy 279 KSLDSGAGCTVAEEVYSSIRNATAGTQDLAKQYEVHLEDEAKWGTNRQIYMGFMIGA 338
Db 225 RITRIYRAAVLAEEVLSIRTVAFYAAQKNVEKYDVFLQAHQGGKKSPNNVGLFST 284
Qy 339 MEGIMTSNGLGFMGMSRFLVDGAV--DVGDILTVLMAIILGSSLSLGNVSPNMAFTNAYA 397
Db 285 EYICITAAIALAFWKCFRMTQNGEVADVGVFTVVLSTVLAASISMLAPSGSVYQRI 344
Qy 398 AAKIFGTIDROSPLDPYSNEGTLDHFRGHTLRLVKNHITPRPVTWMEEDVLSMPAG 457
Db 345 FGSELSIIDKPTQDPLDPDSGQKQPGCGLQIEIQMLAFYRPAQVLRDNLITPAG 404
Qy 458 KTTALVPGSSGSGSTVGVLEFRYMPKRGTVLLDGDIDDLNRWLRQOISVSOEPLV 517
Db 405 KTTALVAGASGSGSTVGVLEFRYMPKRGTVLLDGDIDDLNRWLRQOISVSOEPLV 517
Qy 518 GTTIYKIRHGLGTYEENESEDKVRELLENAMKANANDETALPEGETVNGOGFLL 577
Db 465 RGTIPQNIANGFDEQDQDLPRE--KOMELVQKACKASNGDVFIENLPGYETVEGERAGAL 523
Qy 578 SGGQKORIALARAIVSDPKILLDEATSLDPTKSEGVQVQALERAEGRTTIVIAHRLST 637
Db 524 SGGQKORIALARAIVSDPKILLDEATSLDPAEKVVOEALNRVSKDITIVIAHRLST 583
Qy 638 IKTANHIVLVNCKIAEGOTHELDVDRGAYRKRLVEAQRIT--NEQKADALEDAEDLT 695
Db 584 YKSAGNIAVISOGLIVEQTHHELIEFGCHYALVAQDLGADQEQEHB----- 632
Qy 696 NMDIAKIKTA--SSASDDLCKPTTIDRTGTHKSVSASAILSKAR--PETTKYSLMTLTKV 753
Db 633 -----KTLHEKAKAREAAGERPALERTHTTATSQGDLEKRRVPGTGLYSILKCI-LI 684
Qy 754 ASFNRPETIPY--MLIGLVSVLAGGCGPTQAVLYAKAISTSLPESQYSKLRHDAEFWSLM 812
Db 685 MEYEQKNLWCFLLSTIYVILICATPPGALLFSRLTYFEL---SGHAAQERADPTIIM 741
Qy 813 FRYVGIQFTOSTNGAFAVCSERLIRARSTAFRTILRODIAFEDKENSGLTSLFL 872
Db 742 FFVVALGNIYGFITGWTGNISQVTVHRYQAMFORVADQIDELLDPEQSLTSQL 801
Qy 873 STETKHLGSGSVTLGLTILMTSTLGAIIITIALAIGKMLAVCISVVPVLLACG----- 926

```

```

Db 802 SALPTOLQEL-----ISANPLIYVVGQHLQEC-STSLMNETGPGGCW 846
QY 927 ---FYREYMLAQFOSRS-----KLAVEGSANFACEATS---SIRTVASLTRERDWEIYH 975
Db 847 CTSTPAFGWLPQNSNDDEARSRLGKRCRCMKCRSSYDDPRLRIDSRPCSPYLG 906
QY 976 AOIDAGQRTSLISLVANSSLL-----YASSQALVFVVALGF-----W-YGCTLIGHHE 1032
Db 907 VE---OG---LAKTIOSFMWGRCEFHLSQSMFLAIALGFCIANDWMLQVSTQLN--- 957
QY 1023 YDIFREFVCFRS-ELIFGAOSAGTVESFAPDMGKAKNAAEFRRLFDKRPQIDNMSEGEK 1081
Db 958 -----PISSMAFCLPVOAAQIYLAISTFTAKRSANNTILMLRTLKPIRE-TEENKK 1010
QY 1082 LETVEG-ELIEFRNVHRYPRPEOPVLRGDLTVKPGQYVALVSPGCGSTTIALERE 1140
Db 1011 KGVVGGCPVDLEDIEFRYRQDSARVLRGVSMTIEPGFAVYAGASCGSTIALSERF 1070
QY 1141 YDAIAGSILVDGDKISKLINSTRSFLSVSOEPTLYOGTIKENILIGIVEDDVEEFLI 1200
Db 1071 YDPTSGRISFAHNIEMSPRLYRGHMSILVQOEPTLYQGSVREWYIAL-EAEISEE-LC 1128
QY 1201 KAKKDANIYFIMS--LPEGFNTVVSCKGMLSGGOKORVAIARALLRDKRIILLDEATS 1258
Db 1129 QGRLPAPMLWLTSLYQAKRLAORGMQFSGGOKORUAIARALLRNRKLLILDEATS 1188
QY 1259 ALDSESEKVVQALDAARGRRTIAVAHRLSTIOKADVIYVFDQKIVESGTHSELYOKK 1318
Db 1189 ALDTQSERLVQALDDEASTRTITIAVAHRLSTIRNDVIYIVFANGRIAEIQTIAELQRLR 1248
QY 1319 GRYYELVNLQSL 1330
Db 1249 GRYYEMCLAQSL 1260

RESULT 15
US-09-873-409-2
; Sequence 2, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-409-2

```

Query Match 25.2%; Score 1690.5; DB 10; Length 812;
 Best Local Similarity 43.5%; Pred. No. 1,2e-117;
 Matches 366; Conservative 152; Mismatches 293; Indels 31; Gaps 7;

```

QY 489 LLDGHIKDILNMLRQQLSVSOEPTLYOGTIKENILIGIVEDDVEEFLIKACKDANI 548
Db 1 MYDENDIRALNVRHRYDHIGVSOEPTLYOGTIKENILIGIVEDDVEEFLIKACKDANI 51
QY 549 AAKMANAHDEFTALPEGETNNGORGLISGGOKORAIARAVYSDPKIILLDEATSALD 608
Db 52 AAREANAYDIMEFPKNTLVGEGKQMGOKORAIARALRNPKIILLDEATSALD 111
QY 609 TKSEGVQALAEBAERGRTTIVIAHRLSTIKTANIVLVNKGIAEQGTHDELVDREGAY 668
Db 112 SESKSANVQALAEKASGRRTIIVAHRLSTIRSDLIYTLKDMIAEKGAHAELMAKRGLY 171
QY 669 KRLVEQRITVEQKADALEDDADNADILNADIATKTAASSASDLDGKPTTIIDRTGTHKSV 728

```

```

Db 172 YSLVMSODIKKADQEMESMTYSTERKTNS--LPLHVSXIKSDF-----IDK--AEEST 221
QY 729 SSALISKRPEPTTPKYSLMTLLKRFVASENRPREIYMLIGLVEFVLAGGQPTQAVLAKA 788
Db 222 QSKRIS-----LPEVSLIKLK----LNKPEMFVVLGILASVLNKTVPVSIITRAKI 271
QY 789 ISLTLSESOYSKLRHDAFWSLMFVVGIIQFTOSTNGAFAFVCSERLIRARSTAFR 848
Db 272 ITMGF--NNDKRTLLKHADEIYSMIFVLGVICFYSYFMQGLFERYRAGEIILMLRHLAFK 329
QY 849 TILRODIAFDKENSNGALTSFSTETKHLISGVSYTLGITLMTSTTLGAAILTALAIG 908
Db 330 AMYQODIAWDEKENSNGGLTTLAIDIAQOGATGSRIGVLTQANANMGSLVSIISFIYG 389
QY 909 WKLALVCSIVPVYLACGFREFYMLAQFOSRSKLATYAGSANFACEATSIRTVASLTRER 968
Db 390 WEMFLLISTAPVLAVGMETLTAAMTGFANKDKELKHAGIATAELENITITVSLTREK 449
QY 969 DWELIYHAQIDAQRTSLISVLRSSLLYASSQALVFPCVALGFYVGGTLGHHEDYIFRF 1028
Db 450 AFEQMEEMLOTOHRNTSKNAQIIGSCYAFSHAFYFAVYAGFRFGAYLLIQAGHMTPEGM 509
QY 1029 FVCSSEILFQAQSGIVTFSPAPDMGKAKNAAEFRRLFDKRPQIDNMSEGEKLETYEGE 1088
Db 510 FIYFTALAYGAMALGKTLVLAPEYSKAKSGAHLFALLEKKPNIDSRSOEGKKPDTECGN 569
QY 1089 IEFENVHFRYPTRPEOPVLRGDLTVKPGQYVALVSPGCGSTTIALLERFYDAIGSI 1148
Db 570 LEFEVSEFTPCRPDVTILKGLSLSTERKIVAVAGSSGCKSTYVOLLOKLDYPPVQGV 629
QY 1149 LVGDKDISKLINSTRSFLSVSOEPTLYOGTIKENILIGIVEDDVEEFLIKACKDANI 1208
Db 630 LFDGVDAKELNVQWLBRSQIAIVPEVPLENCSIENIAYGDNVSVPVLDDEIKKANANANI 689
QY 1209 YDFIMSLPEGFNTVVSCKGMLSGGOKORVAIARALLRDKRIILLDEATSALDSESEKVV 1268
Db 690 HSFTEGLPEKYNNQVGLKQALSGGOKORUAIARALLQPKIILLDEATSALDNDSEKVV 749
QY 1269 QALDAARGRRTIAVAHRLSTIOKADVIYVFDQKIVESGTHSELYOKKGRYYELVNLQ 1328
Db 750 QHALDKARTGRTCLVYTHRLSAIONADLIYVLHNGKIKEGOSTHOLLNRRNIYIKLVNAQ 809
QY 1329 SL 1330
Db 810 SV 811

```

Search completed: April 1, 2003, 16:08:22
 Job time : 34 secs

xx The present sequence represents a multiple drug resistance protein, designated atdD, which is isolated from *Aspergillus nidulans*. The atdD polynucleotide is used to produce nucleic acid molecules which are useful in diagnostic assays for the detection of atdD polynucleotides. The atdD polypeptide may be used to identify agents which inhibit multiple drug resistance. These agents may be useful for the treatment of fungal diseases. New antifungal compounds may also be identified using the gene disruption or gene replacement strains of *Aspergillus nidulans*.

xx Sequence 1334 AA:

Query Match 100.0%; Score 6721; DB 20; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSPLEINPSPETAMREPASTSTTEBOASTPHAADEKILISDLSAPSSSTTATPADEENRP 60
DB 1 MSPLEINPSPETAMREPASTSTTEBOASTPHAADEKILISDLSAPSSSTTATPADEENRP 60
QY 61 KSSSSNNAVSEVDALIAHLEPDERQVLKQLEIKVNISEFGIMRYATKMDILMWIS 120
DB 61 KSSSSNNAVSEVDALIAHLEPDERQVLKQLEIKVNISEFGIMRYATKMDILMWIS 120
QY 121 TTTCAATAASTFORIMLYQISTDEYDELTKNLYLYLIGIGERYVYVSTVGFIYTGERRA 180
DB 121 TTTCAATAASTFORIMLYQISTDEYDELTKNLYLYLIGIGERYVYVSTVGFIYTGERRA 180
QY 181 TOKIREYTESILRONIGFEDKIGAGEVTTTITADTNLIQDGISEKVGITLALATFYVA 240
DB 181 TOKIREYTESILRONIGFEDKIGAGEVTTTITADTNLIQDGISEKVGITLALATFYVA 240
QY 241 FTIAVYKWKALICSSITVALVLTMGSGSOFIYKSKSLDSYGAGTVAEVISSIRN 300
DB 241 FTIAVYKWKALICSSITVALVLTMGSGSOFIYKSKSLDSYGAGTVAEVISSIRN 300
QY 301 AFAFGTODKLAQOYEYHDEAEKNGKNOIYNGFMIGAMFGIMSYNGLGFMWGSFELVD 360
DB 301 AFAFGTODKLAQOYEYHDEAEKNGKNOIYNGFMIGAMFGIMSYNGLGFMWGSFELVD 360
QY 361 GAVDVGDIITVLMALIGSFSIGSNVSPNAQFTNAVAANAATFGTIDRQSPIDPYSNECK 420
DB 361 GAVDVGDIITVLMALIGSFSIGSNVSPNAQFTNAVAANAATFGTIDRQSPIDPYSNECK 420
QY 421 TLDHFEGHIELRNVKHIYPSREPVYMEDVSLSPAGKTTALVGPSSGSKSTVVGLEVERF 480
DB 421 TLDHFEGHIELRNVKHIYPSREPVYMEDVSLSPAGKTTALVGPSSGSKSTVVGLEVERF 480
QY 481 YMPVGGTVLLDGHDKDLRWLRQOISVSOEPVLFGTITIKNIRHGLIGTRYENESHD 540
DB 481 YMPVGGTVLLDGHDKDLRWLRQOISVSOEPVLFGTITIKNIRHGLIGTRYENESHD 540
QY 541 KYRELIENAKKANAHDFITALEGETVNGQGFLLSGGQKORIAIARAVVSDPKILL 600
DB 541 KYRELIENAKKANAHDFITALEGETVNGQGFLLSGGQKORIAIARAVVSDPKILL 600
QY 601 DEATSAIDTKSEGVQALERAEGRTTIVIAHRLSTIKTANHIVLVNGKIAEOGTDE 660
DB 601 DEATSAIDTKSEGVQALERAEGRTTIVIAHRLSTIKTANHIVLVNGKIAEOGTDE 660
QY 661 LVDRGAVYKLVLEAQRINQKQADALEADADLTNADIAKIKTASASSDLDGKPTTID 720
DB 661 LVDRGAVYKLVLEAQRINQKQADALEADADLTNADIAKIKTASASSDLDGKPTTID 720
QY 721 RTGTAKSVSSAITSKRPPETPKYSIMTLTKFVASFNREPIYMLIGLVSYLAGGGQPT 780
DB 721 RTGTAKSVSSAITSKRPPETPKYSIMTLTKFVASFNREPIYMLIGLVSYLAGGGQPT 780
QY 781 GAVLYAKAISTSLPESQYSLRHADAFWLSMFVVGIIQFTQSTNGAFAVCSERLIR 840
DB 781 GAVLYAKAISTSLPESQYSLRHADAFWLSMFVVGIIQFTQSTNGAFAVCSERLIR 840

```

```

QY 841 RARSTAFRTILKODIAFPDKEENSTGALTSFSTERKHLISGSGVTLGTLMTSTTLGAA 900
DB 841 RARSTAFRTILKODIAFPDKEENSTGALTSFSTERKHLISGSGVTLGTLMTSTTLGAA 900
QY 901 IITAIALIGKIALVLCISVVPVLACGFYRFYMAOFOSRKAYAGSANAFCEATSTIRT 960
DB 901 IITAIALIGKIALVLCISVVPVLACGFYRFYMAOFOSRKAYAGSANAFCEATSTIRT 960
QY 961 VASLFRERDWEIYHAQDLDAQGRSLISVLRSSLLYASSQALVFQVALGFVYGGTLLGH 1020
DB 961 VASLFRERDWEIYHAQDLDAQGRSLISVLRSSLLYASSQALVFQVALGFVYGGTLLGH 1020
QY 1021 HEYDIFRFVFCSEILIFGASQGTFSFAPDMGKAKNAAEFRRLPDRKPOIDNNSEGE 1080
DB 1021 HEYDIFRFVFCSEILIFGASQGTFSFAPDMGKAKNAAEFRRLPDRKPOIDNNSEGE 1080
QY 1081 KLETVGEIEFRNVHFRYTRPEQPYLRGLDLTVRPGQVVALVGPSCGKSTTIALLERF 1140
DB 1081 KLETVGEIEFRNVHFRYTRPEQPYLRGLDLTVRPGQVVALVGPSCGKSTTIALLERF 1140
QY 1141 YDAIAGSILVDSKDISKLNINSYRSFSLVSOEPTLYOGTIRENILLGIVEDVPEEFLI 1200
DB 1141 YDAIAGSILVDSKDISKLNINSYRSFSLVSOEPTLYOGTIRENILLGIVEDVPEEFLI 1200
QY 1201 KACKQANITYDFIMSLPEGRNTVYSGKGMLSGGQKQRAVIAAALLRDPKILLDBATSL 1260
DB 1201 KACKQANITYDFIMSLPEGRNTVYSGKGMLSGGQKQRAVIAAALLRDPKILLDBATSL 1260
QY 1261 DSESEKVOAALDAARAGRTTIAVAHRLSTIQKADVIYFDPDGKIVESGTHSELVOKGR 1320
DB 1261 DSESEKVOAALDAARAGRTTIAVAHRLSTIQKADVIYFDPDGKIVESGTHSELVOKGR 1320
QY 1321 YYELVNLQSLGKH 1334
DB 1321 YYELVNLQSLGKH 1334

```

RESULT 2
AA16434
ID AA16434 standard; Protein; 1349 AA.
XX
AC AA16434;
XX
DT 10-AUG-1999 (first entry)
XX
DE Multiple drug resistance protein of *A. fumigatus*.
XX
KW Multiple drug resistance protein; AfumDRL;
XX multiple drug inhibition activity.
XX
OS *Aspergillus fumigatus*.
XX
PN US914246-A.
XX
PD 22-JUN-1999.
XX
PF 08-MAR-1996; 96US-0612734.
XX
PR 08-MAR-1996; 96US-0612734.
XX
PA (ELI) LILLY & CO ELI.
XX
PI Peery RB, Skatrud PL, Tobin MB;
XX
DR WPI; 1999-370499/31.
XX
PS N-Psdb; AAX60201, AAX60202.
XX
PT Drug resistance gene from *Aspergillus fumigatus* useful for
XX determining fungal multiple resistance inhibition activity
XX
PS Claim 3; Columns 21-28; 25pp; English.
XX
CC The present sequence represents the multiple drug resistance protein of

CC Aspergillus fumigatus, designated AfumDR1. The specification describes
 CC a method whereby cultured transformed fungal cells are used to
 CC determine the multiple drug inhibition activity of a compound.
 CC The method is useful for providing nucleic acid molecules which
 CC encode multiple drug resistance proteins from Aspergillus fumigatus
 CC (Afuf) which are useful for determining the MDR inhibition of a compound.
 CC XX

Sequence 1349 AA;

Query Match 80.5%; Score 5407.5; DB 20; Length 1349;
 Best Local Similarity 79.1%; Pred. No. 0;
 Matches 1069; Conservative 114; Mismatches 141; Indels 27; Gaps 6;

QY 5 EFNPLSPETAMREPASTTEBOASTPHAADEKKILSDISAPSSYATPADKRRPKSSS 64
 DB 5 EFGASSREKSL-EDLQVATLEKGRSTSSSGADNEKPHDHSLSDTIMAPPD---GKKKD 59
 QY 65 SNNASVNEVDALIAHLPEDEROYLTOLEIKVINISFEGLMYARKMDILINIVISTICA 124
 DB 60 HGRKAVDLND-DSLFAHLOHEKEVLRKOLDAPSVKVSFPLYRYASRKDILILVSAICA 118
 QY 125 IAA-----ASTFORIMLOISYDEYDELFRKNLYEYVYLGIGEEVTVYST 170
 DB 119 IAAGAALPLFTILFGSLASFQISLCTMYHEFYHKLTKNVLFEYVYLGIAEFVTVYST 178
 QY 171 VGEITYGEHATOKIRREYVESILRONIGYFDKLAGEVTTTRITADYNLIDGISEKVLG 230
 DB 179 VGEITYGEHATOKIRREYVESILRONIGYFDKLAGEVTTTRITADYNLIDGISEKVLG 238
 QY 231 LVALATPFAFIAYVYKYNKALIGSSTVAVLVNMGSGSOPFIKYSKSLDSYGAGTV 290
 DB 239 LVAFAFVFAFIAYVYKYNKALIGSSTVAVLVNMGSGSOPFIKYSKSLDSYGAGTV 298
 QY 291 AEEVISINNAFAFGTODKLAKOYEVLDEAEKMGTKNOYVGMFGMIGMGLWYNSYGLG 350
 DB 299 AEEVISINNAFAFGTODKLAKOYEVLDEAEKMGTKNOYVGMFGMIGMGLWYNSYGLG 358
 QY 351 FWMGSRFLVDGAVDGLITLVMAILLIGSFLGNSPNAQFTNNAVAANAIRIGTIDROS 410
 DB 359 FWMGSRFLVDGAVDGLITLVMAILLIGSFLGNSPNAQFTNNAVAANAIRIGTIDROS 418
 QY 411 PLDPYSNECKTLDHFEHGHLELRNVKHTYPSRPVYMEVYSLSPAGKTATVAPSGSGK 470
 DB 419 PLDPYSNECKTLDHFEHGHLELRNVKHTYPSRPVYMEVYSLSPAGKTATVAPSGSGK 478
 QY 471 STYVGLVERFYMPYRGTVLLDGHDKDLNLRMOOISLVSQEPVLEFGTTTYKNIRHGLI 530
 DB 479 STYVGLVERFYMPYRGTVLLDGHDKDLNLRMOOISLVSQEPVLEFGTTTYKNIRHGLI 538
 QY 531 GTRVNESEDKVLELEENAKMANAHDFITALPEGYETNVGQGFLLSGOKORAIARA 590
 DB 539 GTRVNESEDKVLELEENAKMANAHDFITALPEGYETNVGQGFLLSGOKORAIARA 598
 QY 591 VVSDPKILLDEATSLDTRKSEGVVQALERAAGRTTIVIAHRLSTIKTAHNIIVLVNG 650
 DB 599 VVSDPKILLDEATSLDTRKSEGVVQALERAAGRTTIVIAHRLSTIKTAHNIIVLVNG 658
 QY 651 KTAEOGTDELVDGAYRKLVEAORINQOKDALE---DADAPELTNADIKKITASS 707
 DB 659 KTAEOGTDELVDGAYRKLVEAORINQOKDALE---DADAPELTNADIKKITASS 718
 QY 708 ASSDLG---KPTTIDRTGTHKSVSAILSKRPETTPKYSIMTLILKFAVSFNRPEIPY 763
 DB 719 SSNSLDAVEDEKALEMKRGTKGTSVSAVLSKVPQEFKYSIMTLILKFAVSFNRPEIPY 778
 QY 764 MLIGLVSVLAGGGPTQAVLVYAKAISTLSLPSQYSKLRHDDFWSLMFVVGITQFIT 823
 DB 779 MLIGLVSVLAGGGPTQAVLVYAKAISTLSLPSQYSKLRHDDFWSLMFVVGITQFIT 838
 QY 824 QSTNGAFAFVCSERLIRRASTAFPTILRODIAFPDKEENSTGALTSFSTETKHISGVS 883
 DB 839 LSTNGAFAFVCSERLIRRASTAFPTILRODIAFPDKEENSTGALTSFSTETKHISGVS 898

QY 884 GVTGLTLMSTMTLGAIIITAIAGMKLALVCISVVPVLLACGRYPRFMYLDAQOSRSKLA 943
 DB 899 GVTGLTLMSTMTLGAIIITAIAGMKLALVCISVVPVLLACGRYPRFMYLDAQOSRSKLA 958
 QY 944 YEGSANACATSSIRFVASTREBDVETLYHMDLQNGSTSLISVRSLLTVASSQALV 1003
 DB 959 YEGSANACATSSIRFVASTREBDVETLYHMDLQNGSTSLISVRSLLTVASSQALV 1018
 QY 1004 FFCVVALGFMYGTLTGHHEEDIFFEFVCFSEILFGAOSAGTVSEFADPMKAKNAAEFR 1063
 DB 1019 FFCVVALGFMYGTLTGHHEEDIFFEFVCFSEILFGAOSAGTVSEFADPMKAKNAAEFR 1078
 QY 1064 RLFDKRPQIDNWSSEKLETVEGEIEFRVNHRYRYPREPQVYLRGLDITVKKQYVALV 1123
 DB 1079 RLFDKRPQIDNWSSEKLETVEGEIEFRVNHRYRYPREPQVYLRGLDITVKKQYVALV 1138
 QY 1124 GPSCGCKSTTIALLERVYDALAGSILVDGDISKLNINSTRSFLSLYSQEPITYQGTIKE 1183
 DB 1139 GPSCGCKSTTIALLERVYDALAGSILVDGDISKLNINSTRSFLSLYSQEPITYQGTIKE 1198
 QY 1184 NILGIYEDVPEBEFLKACKDANIYDFIMSLEPGFNTVVGSGMGLSGGOKORVALARA 1243
 DB 1199 NILGIYEDVPEBEFLKACKDANIYDFIMSLEPGFNTVVGSGMGLSGGOKORVALARA 1258
 QY 1244 LTRDPKILLDEATSLDSESEKVVQALDAAARGRTTIVAHRLSTTORADVITYVEDOG 1303
 DB 1259 LTRDPKILLDEATSLDSESEKVVQALDAAARGRTTIVAHRLSTTORADVITYVEDOG 1318
 QY 1304 KIVSGTHSELVQKGRYELVNLQSLGKH 1334
 DB 1319 KIVSGTHSELVQKGRYELVNLQSLGKH 1349

RESULT 3
 AAB62495
 ID AAB62495 standard; Protein; 1333 AA.
 AC AAB62495;
 XX
 DT 09-JUL-2001 (first entry)
 XX
 DE P. chrysogenum ABC transporter a10 sequence.
 XX
 KW Beta-lactam; micro-organism; ABC transporter; Amp-binding cassette;
 KW cephalosporin; a10.
 XX
 OS Penicillium chrysogenum.
 XX
 PN MO200132904-A2.
 PD 10-MAY-2001.
 XX
 PF 03-NOV-2000; 2000MO-EP11489.
 XX
 PR 03-NOV-1999; 99EP-0203684.
 PR 03-NOV-1999; 99EP-0203685.
 PR 03-NOV-1999; 99EP-0203687.
 PR 03-NOV-1999; 99EP-0203688.
 PR 03-NOV-1999; 99EP-0203689.
 PR 03-NOV-1999; 99EP-0203690.
 PR 03-NOV-1999; 99EP-0203691.
 PR 03-NOV-1999; 99EP-0203692.
 PR 03-NOV-1999; 99EP-0203693.
 PR 03-NOV-1999; 99EP-0203694.
 XX
 PA (STAM) DSM NV.
 PI Van Den Berg MA, Bovenberg RAJ, Driessen AJM, Konings WN;
 PI Schuurs TA, Nieboer M, Westerlaken I;
 XX WPI; 2001-291055/30.
 DR N-PDB; AAF83399, AAF83400.

xx Enhancing secretion of beta-lactam compounds from a micro-organism by
 PT enhancing adenosine triphosphate-binding cassette transporter activity,
 PT useful for producing e.g. penicillin and cephalosporins -
 xx
 PS Claim 8; Page 110-112; 116pp; English.
 CC
 CC The invention relates to a method for enhancing the secretion of beta-
 CC lactam compounds from a micro-organism that comprises enhancing ABC
 CC (adenosine triphosphate (ATP)-binding cassette) transporter activity of
 CC the micro-organism. The method is used for enhancing the production and
 CC secretion of beta-lactam antibiotics like penicillin and cephalosporins.
 CC Manipulation of the ABC-transport protein system provides a means for
 CC enhancing beta-lactam secretion. The present sequence represents the
 CC P. chrysogenum ABC transporter aat0 protein sequence.
 CC
 xx Sequence 1333 AA;
 SQ
 Query Match 76.1%; Score 5117.5; DB 22; Length 1333;
 Best Local Similarity 75.7%; Pred. No. 0;
 Matches 1015; Conservative 135; Mismatches 158; Indels 33; Gaps 7;
 15 MPEPSTTEQASTPHADEKKT-LSLSAPSSSTTAPADKEHRSKSSNNNAVNE 73
 1 MOQPEGVAVKADGPDQNKPVDRKSDASTSTA---KEHRSKTPVA-----SD 50
 74 VDALLAHLPEDEROYLKTQLEIKYNIISFGLMRATKMDILMYISTCAIAA----- 127
 51 EDALAHLPHEHKDLKQDLAPVNIISFGLYRASRDILIIYSTCAIAAGALPL 110
 128 -----ASTQRIMLYOISYDERYDELTKNVLYFVYIGIEFVTVVYVIGFIYTGSH 179
 111 FTILGSLATAPQKIMLRPIPEDEYDQLTVMVLYFYIGIEFVTVVYVIGFIYTGSH 170
 180 ATOKIREYIESILNONIGYFDPKLGAEYTRITADTNLIDGISEKVELTTLATETPT 239
 171 VTOKIREHLEILMONIAYFRLKLGAEYTRITADTNLIDGISEKVELTTLATETPT 230
 240 AFIAYVYKWLALICSSSTVALVLMGGGSOFIIRKSKSDSYAGAGTVAEEVSSIR 299
 231 AFIYAYIKYAPLAGICTSTMVALVLMGGGSOIIRKSKSDSYAGAGTVAEEVSSIR 290
 300 NATAGTODKLAKOYEHLDEAEKWTGNQIYMGFMICAMGELMYSNGLGTFMGSRLY 359
 291 NATAGTODKLAKOYESHLLRAERWGMRLQMSLAVVGMFIMMNYGLGTFMGSQFLY 350
 360 DGAVVVDGLTVLMAILIGSFSLGNVSPNAOFTNVAANAIFGTIDQSPLDPSNMG 419
 351 DGKVVYGVYLTLMAILIGSFSLGNVSPNAAFNTNVAATRIKATIDQSPLDPSDGG 410
 420 KTLDFEGHIELRNVKHIYPSRPEYTVMEDVSLSPAGKTTALVPSGSGSTVGLVER 479
 411 ILDFHEGHIEFRNVKHIYPSRPEYTVMDVSLGIPAGNTTALVPSGSGSTVGLVER 470
 480 FYMPVRCGYLDDGHDKDLNLMWLNOQISLVSEPEYLVSTTYKINRHLIGTKYENEE 539
 471 FYLPVGGVYFDGHDIGQTNLMRLNOQISLVSEPEYLVSTTYKINRHLIGTKYENEE 530
 540 DKVRELIENAKMANAHDFTALPEGYETNVGREGFLISGGOKORAIARAVVSPKILL 599
 531 EKIKELIENAKMANAHDFTALPEGYETNVGREGFLISGGOKORAIARAVVSPKILL 590
 591 IDEATSAIDTKSEGVVQALDRAAGRTTIVAHRLSTIKSHNIVYFVQSGIVEGTHS 650
 600 IDEATSAIDTKSEGVVQALDRAAGRTTIVAHRLSTIKSHNIVYFVQSGIVEGTHS 659
 660 ELYVDGGAIRKLYEAKRINEKEADAL---EDADA-EDLTNADIKIKTASSASDGLVPS 715
 651 QLTEDHGPYFKLYEAKRINEKEADALDADEDEGLIEETKSHIRAVKSIAGSGTCVKE 710
 716 PTT---IDRTGTHSVSAIISKRPETTPKYSIMTLTKFVASFNRPPIPYMLGLVPS 771
 711 AETPDAMHROGSRKSVSSVILSOKTAEGRKHSLLTLIKFISFNKEERWFMALGLCHS 770

QY 772 VLAGGQPTQAVLYAKAISTLSLPSOYSKLRRHADFMELAFVYGIOTFQSTNGCAF 831
 DB 771 ILAGCGPTQAVLYAKAISTLSLPSOYSKLRRHADFMELAFVYGIOTFQSTNGCAF 830
 QY 832 AVCSERILRRARSTAFRTLLRODIAFPDEKENSTGALNLSLSTERKHLGSGVLTGTL 891
 DB 831 AFCSERILRRARSTAFRTLLRODIAFPDEKENSTGALNLSLSTERKHLGSGVLTGTL 890
 QY 892 MNTSTLGAIIITAILMGWLTALVCSIVPVYLLACGYPRTYMLAQOSRKLAYESSANPA 951
 DB 891 MNTSTLGAIIITAILMGWLTALVCSIVPVYLLACGYPRTYMLAQOSRKLAYESSANPA 950
 QY 952 CEATSIIRTVASLTREEDVWEIYHQAOLDAGRTSLISVRSLLIYASSQALVFCVALGF 1011
 DB 951 CEATSIIRTVASLTREEDVWEIYHQAOLDAGRTSLISVRSLLIYASSQALVFCVALGF 1010
 QY 1012 WTGFTLLGHEHYDIFRFVCFSEILFGAQSAGTVSPFADPMKAKNAAEFRLDRKPO 1071
 DB 1011 WTGFTLLGHEHYDIFRFVCFSEILFGAQSAGTVSPFADPMKAKNAAEFRLDRKPO 1070
 QY 1072 IDNWESEGEKLETVGEIEFRVVRHRYPTRPPOVLRGLDLYKRGQYALVGPBGCGKS 1131
 DB 1071 IDNWESEGEKLETVGEIEFRVVRHRYPTRPPOVLRGLDLYKRGQYALVGPBGCGKS 1130
 QY 1132 TTITALERPYDAIAGSILVDGDISKLNINSYRSEFLSIVSQPPTYOGTIRKENIILGIVE 1191
 DB 1131 TTITALERPYDAIAGSILVDGDISKLNINSYRSEFLSIVSQPPTYOGTIRKENIILGIVE 1190
 QY 1192 DQVEPEELIKACKDANITDFIMSLEPGENTVYVSGGMISGOKORAVATARALDRPKIL 1251
 DB 1191 DQVEPEELIKACKDANITDFIMSLEPGENTVYVSGGMISGOKORAVATARALDRPKIL 1250
 QY 1252 LDEATSAIDSESEKVVQALDRAAGRTTIVAHRLSTIKSHNIVYFVQSGIVEGTHS 1311
 DB 1251 LDEATSAIDSESEKVVQALDRAAGRTTIVAHRLSTIKSHNIVYFVQSGIVEGTHS 1310
 QY 1312 SELVOKRGRYEVLNLOSLGR 1332
 DB 1311 TELLRKGRYFELVNLOSLGR 1331
 RESULT 4
 ID AAR9255 standard; Protein; 1307 AA.
 XX AAR9255;
 XX
 XX AAR9255;
 XX
 XX 01-DEC-1996 (first entry)
 DE Aspergillus flavus multidrug resistance protein Afl-MDR1.
 XX
 XX Multiple drug-resistance; MDR; Afl-MDR1; antifungal.
 OS Aspergillus flavus.
 XX
 XX W09626952-A1.
 XX
 XX 06-SEP-1996.
 XX
 XX 23-FEB-1996; 96WO-US02665.
 XX
 XX 27-FEB-1995; 95US-0395246.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 XX Peery RB, Skatrud PL;
 PI
 XX WPI; 1996-412733/41.
 DR
 XX N-PSDB; AAT35230.
 XX
 PT New multiple drug resistance gene of Aspergillus flavus - used to
 identify cpds. that inhibit fungal multiple drug resistance activity

xx Disclosure; page 22-27; 37pp; English.
 ps A new multiple drug resistance (MDR) protein (AAR9255), designated
 xx Afl-MDR1, of *Aspergillus flavus* was identified as the product of
 CC a cDNA clone isolated from a mycelial cDNA library of *A. flavus*.
 CC Afl-MDR1 can be synthesised by host cells, partic. *Saccharomyces*
 CC *cerevisiae*, transformed with vectors that provide for the expression
 CC of DNA encoding Afl-MDR1. Transformed host cells can be used to
 CC identify cpds. that inhibit fungal MDR activity.
 xx
 xx Sequence 1307 AA;
 Query Match 59.4%; Score 3992; DB 17; Length 1307;
 Best Local Similarity 60.2%; Pred. No. 0;
 Matches 790; Conservative 211; Mismatches 244; Indels 66; Gaps 11;
 60 PRSSSSNNNAVSEVDALIAHLPEDEROYLKTQLEIKVNISEFGIMRRATKMDILIMVI 119
 13 PKSPGTGSGT-----GHSVSHAEVLDRLHTPVSOIGFGIYRYATRMVAILFG 64
 120 SITCAIAA-----ASTFORIMVOISYDEFEYDELTKVNVYFVYLGIEFVY 165
 65 SALAALAGGALPLFTVLFGRLTSTFQDIATHTTYDHFHHELTKNVYFVYLGAAEFVA 124
 166 VYVSTVGFYTGSHATQIKREYVLESILRONIGYFDFKLAGEVYTRITADTNLIQGISSE 225
 125 IYLATVGFYTGSHVYQIRVEYFOALRNQNTAFEDTLGAGETTRITADTNLIQGISSE 184
 226 KVGILITLALATVYATITAVYKWKALICSSITVALVLTMGSGSOFITIKSKSIDSYG 285
 185 KVGILITLSTVYATITAVYKWKALICSSITVALVLTMGSGSOFITIKSKSIDSYG 244
 286 AGGTVAEEYISSIRNATAGTODKLAKEYEHVHDEAEKGTGKQIYMGFMIGMFGIMYS 345
 245 RGSNAEDILDSIRYAAINAQETLAKTESHDKABEGMKSKVIFALWGLICMVL 304
 346 NYGLGFMWGSREFLVDA--VDVGDILTVLMAILIGSESLGNVSPNAQFTNAVAAAATIP 403
 305 NYGLGFMWGSREFLVESINIKAGDVLTIMAAITIGSYNIGNVAPNQAALSDVAAAASKLV 364
 404 GTIDRQSPIDPYSNECKTLDHFEGLHLENVKNHITYPSREPVYMEDVSLSPAGKTTALV 463
 365 GTIDRQSPIDALSDQCKTEFVGNVIVLQNIKRVYPSREPVYVAHDLSCITIPAGKTTAFV 424
 464 GPGSGKSTVVGLEVERFYMPVREGTVLLDGHIDNLRLRQOISLYSOEVPFGTTIK 523
 425 GPGSGKSTTISLERFYDPVAGTIMGDHDIOQLMLRLRQOISLYSOEVPFGTTIAE 484
 524 NTRHGLIGTYENESDKEYRELIENAKMANADFTALPEGEYETVNGORGFLLSGGOK 583
 485 NTRYGIGSRFEKESYEYELRKRYEAARANADFTALPENGVDINI--ESFSLSGGOK 542
 584 RAIARAAYVSDPKILLDEATSAIDTKSEGVQVQALERAAGRTTYIAHRLSTIKTAHN 643
 543 RAIARAIIITKDPILLDEATSAIDTKSEKLVQVQALDKASKGRTTYIAHRLSTIKAYN 602
 644 IYVLVNGKTAEGSTHDEIVDRGAYRKIVQORINQEKADALDADAEDLTN----- 666
 603 IYVLVNGKTAEGSTHDEIVDRGAYRKIVQORINQEKADALDADAEDLTN----- 660
 697 -----ADIATIKTAASSASDLGKPTTIDRTGTR-----KSSVAILSK-R 736
 661 MTEFFPKDYPGDESDIYSI--LSDASDI-----GLHTGEKORVSRMSLSHLM 708
 737 PPEPTTKYSLMTILKFAVASFNRPEIYMLIGLVSVLAGGGOTQVAVLAKAISTISLPE 796
 709 QPVEKAYEFMTILKFLASNNRPEMPFLLLGLCASILAGIOPSQVAVLFAKAAVSTISLP 768
 797 SOTSRLRHADFWSLMEFVVGIIQFIOSTNGAFAVCSRLRLRAASTAFRTILRODIA 856
 769 LEYPKLRHANPWCMLFMIGIVSLVYISVQGLTFAYSSKMYRARSOAFRVILRODIS 828

QY 857 FPDKENSTGALNLSFSTETKHLSGVSGVLTGLTMTSTTIGAAIIYALIGKRLALVCI 916
 DB 829 FPDQENTGALNLSFSTETKHLSGVSGVLTGLTMTSTTIGAAIIYALIGKRLALVCI 888
 QY 917 SVYVLLACGFYFFYMLAQOSKSLAYEGSANFACATISITVYASLTREDDWELIYA 976
 DB 889 SAVALMCMCFVVMMLERFORAKKAYQESASACBAASAIRTVSLTMTETALOSYQA 948
 QY 977 QLAOGSTLSISLRSSSLVASSOALVEFCVAGLGFVWGTLGHHEHDIRFPVCFSEIT 1036
 DB 949 QLRRLKSLDPLTVKSSLSLASSOALVEFCVAGLGFVWGTLGHHEHDIRFPVCFSEI 1008
 QY 1037 FGAQSAQTVSFAPDMGKAKNAAEFRRLDRPKPOLDMWSESEKLETVGEIEFRVNF 1096
 DB 1009 FGAQSAQTVSFAPDMGKAKNAAEFRRLDRPKPOLDMWSESEKLETVGEIEFRVNF 1067
 QY 1097 RYTPRPQVPLRGIDLTIVKQGYVALVBPBGCGKSTTIALLEFYDAIAGSILYDGDIS 1156
 DB 1068 RYTPRPQVPLRGIDLTIVKQGYVALVBPBGCGKSTTIALLEFYDAIAGSILYDGDIS 1127
 QY 1157 KLINISYRSLVSOEPTLYOGTIRENILLGIYEDVPEPEPLIKCKDKANIDETMSLP 1216
 DB 1128 TLEMSSTRSHLALISOEPTLYOGTIRENILLGIYEDVPEPEPLIKCKDKANIDETMSLP 1187
 QY 1217 EGFNTVVGSKGMLSGGOKORVIAARALDRPKILLDEATSAIDSESEKRVQAAALDAA 1276
 DB 1188 OGFNTVVGSKGMLSGGOKORVIAARALDRPKILLDEATSAIDSESEKRVQAAALDAA 1247
 QY 1277 RGRTTIVAHRLSTIQKADVIYVFDQKIVYSGTHSLVOKKGRYYLVNLOS 1329
 DB 1248 RGRTTIVAHRLSTIQKADVIYVFDQKIVYSGTHSLVOKKGRYYLVNLOS 1300

RESULT 5

AAB62493
 ID AAB62493 standard; Protein; 1261 AA.
 AC AAB62493;
 XX
 XX
 DT 09-JUL-2001 (first entry)
 XX
 DE P. chrysogenum ABC transporter a7 sequence.
 XX
 KW Beta-lactam; micro-organism; ABC transporter; Amp-binding cassette;
 KW adenosine triphosphate-binding cassette; antibiotic; penicillin;
 KW cephalosporin; a7.
 XX
 OS Penicillium chrysogenum.
 XX
 PN MO200132904-A2.
 PD 10-MAY-2001.
 XX
 PE 03-NOV-2000; 2000MO-EP11489.
 XX
 PR 03-NOV-1999; 99EP-0203684.
 PR 03-NOV-1999; 99EP-0203685.
 PR 03-NOV-1999; 99EP-0203687.
 PR 03-NOV-1999; 99EP-0203688.
 PR 03-NOV-1999; 99EP-0203689.
 PR 03-NOV-1999; 99EP-0203690.
 PR 03-NOV-1999; 99EP-0203691.
 PR 03-NOV-1999; 99EP-0203692.
 PR 03-NOV-1999; 99EP-0203693.
 PR 03-NOV-1999; 99EP-0203694.
 XX
 PA (STAM) DSM NV.
 PI Van Den Berg MA, Bovenberg RAJ, Driessen AJM, Konings WN;
 PI Schuurs TA, Nieboer M, Westerlaken I;
 XX WPI; 2001-291055/30.
 DR N-PDB; AAF83393; AAF83394.

XX Enhancing secretion of beta-lactam compounds from a micro-organism by
 PT enhancing adenosine triphosphate-binding cassette transporter activity,
 PT useful for producing e.g. penicillin and cephalosporins -
 XX
 PS Claim 8; Page 92-94; 116pp; English.
 CC The invention relates to a method for enhancing the secretion of beta-
 CC lactam compounds from a micro-organism that comprises enhancing ABC
 CC (adenosine triphosphate (ATP)-binding cassette) transporter activity of
 CC the micro-organism. The method is used for enhancing the production and
 CC secretions of beta-lactam antibiotics like penicillin and cephalosporins.
 CC Manipulation of the ABC-transport protein system provides a means for
 CC enhancing beta-lactam secretion. The present sequence represents the
 CC P. Chrysogenium ABC transporter aa7 protein sequence.
 XX
 SO Sequence 1261 AA;
 Query Match 44.3%; Score 2978; DB 22; Length 1261;
 Best Local Similarity 47.3%; Pred. No. 8.5e-238;
 Matches 603; Conservative 244; Mismatches 383; Indels 44; Gaps 11;
 85 EROVKTQLEIEIKVNSFEGLMRVATKMDILIMVSTICA-IAASTFORIMLYQI--- 139
 6 EOEIIAQOITPIIKINLOEFRTATKQWELIGVSEFLAATITGATITMPALLTGLIGS 65
 140 -----SYDEFYDELTKNVLFVYVGLIGFVTVVYVSGFIYTGHAOKIREYLE 190
 66 IQASWSGSPQDRNSSELTFRITFYVYLFELGIVSCYIANTGIFITGILSRIMERYLA 125
 191 SILRONITFPDKLAGEVTRITADTNLIQDISKVGILTLTALTEFYAFIAYKTKWK 250
 126 ALLSONIFPFNIGAGESTRITTDANLIRGISSEKSVAAVOCASVYAAAFVISEMRMR 185
 251 LALICSTIVLVLMGGSGQFIITYSKSLDSYGAGGVAAEVSIRNATAFGTODXL 310
 186 LTLIASSLICIAVFAAAGIMLTIRQOWLGETRESGITVEEVSIRTVYGLAQSLEL 245
 311 AKOYEHLDEAEKMGTKNQIYVGMIGAMFGIMSYNGIGFWMGSRFLVDAGVADYDILY 370
 246 VARYGSLAKKERFANNRNLISGALLGAVAVIYLAIGFGMGSRLPLVAGSSVYDVL 305
 371 VLMALISFSIGNVSPNAQAFNTAVAAAKITFTIDNQ-SPLDYSNDEGKLDHFECHI 429
 306 IILAVVTGIACIGGIVPLQVFTTAAAGSRLYSTIDRRPSTANRFSBC-SLDSVLGHI 364
 430 ELRANKHLYPSPEVTMEDVSLMPACKTTALVPGSGKSTVVGLEVERFMPRGVYL 489
 365 ELQNRHLYPSRPDIYVDNLSLDIEPKTTAIVPGSGSKSTIELLERFYDPVSGDIL 424
 490 LDGHDIKDLNRLMQQISLVQSQEPVLFCTTYKNIRHGLIGTKYENSEDKVRELIEA 549
 425 LDGHLSELSPMLRQQLISLVQSQPTLPATTFENIRYGLVGTPEMNASRENIENIVGA 484
 550 AKMANAHPTALPBGYETNGQKGFLLSGGOKRIATARAIVSPKILLDEATSAIDT 609
 485 TRLANAHPTIKLPDGYDLVGEAGVLLSGGOKRIATARAISPRILLDEATSAIDS 544
 610 KSEGYYOALAEBAAGRTTYIAHRLSTIKTAHNIVLVNKGIAOGTHDELVDGAYR 669
 545 MSESIVQAHIEKASQRTTYIAHRLSTIKTAHNIVLVNKGIAOGTHDELVDGAYR 604
 670 KLVEAQRINEQKADALDADADLTNADIAKIKTASSASSDLDKPTTIDRTG----TH 725
 605 RLAKQAANVLSER---KRPDDDSLGVPILPAKSVSESE--KAPPLVSGGGLPFDIV 658
 726 KVSVAIIS-KRPETTKYSLMTLKLKTVASFNRPETPYMLIGLVSVLAGGGOPTQAVL 784
 659 RKGAALASLRPPH---QYSWQTLRFVNSHKDSIKLVSGMGLASIQTGAGAVQAVF 715
 785 YAKAISTSLPESQYSKLRHADFWSLMFVYGIQFITQSTNGAFAVCSRLIRRRARS 844
 716 LAKCLVALARPTESPQRLSETNMAGHVVYLAFOFPAYSQAQASALGKCTERILRLGD 775

QY 845 TAFRTILRODIAFPDKEENSTGALNSFLSTETKILSGVSVTLGLTMTSTTLGAIIIA 904
 DB 776 LSFALLLETDMSSFPMEEHGVGALVSLGTETBPASMAHGCAVIGNYIMALTTLGAIAMS 835
 QY 905 LAIGKRLALVCSVYPVLLACGFYRFYMLAQRQSKLAIYESNAFACENRSITVAVSL 964
 DB 836 MAVGKRLISVGAATVPVLLMCGELFRFVMAQLEAHLROYQETASLASEVASAITVLSL 895
 QY 965 TPERVMEIYHAQDAGRTSLISVRSILYASSQALVFCVAGLGFYWGTLGHHEVD 1024
 DB 896 NRESEVAVFHKLKAEQDSKISRSLSSGLFAFQSAPLCTALGLWYGTGLVSGEYG 955
 QY 1025 IFRFVCFSEILFGAQSAGYVFSFAPDMGKAKNAAEFRRLFRRRPQIDMNSSEGEKLET 1084
 DB 956 LFOFILSFAAVNICGDAAGSIRSSPDIAKAKLSTRLKGLLDROPQ-----RPIDF 1007
 QY 1085 VE-----GEIFRNVHRRYPTRPBPQVLRGLDLTVKQGYALVGPSCGKSTIALLE 1138
 DB 1008 VEPATLIGTEIFRNVHRSYTPRPQQLKGLDLTVHKGYALVGPSCGKSTIALLE 1067
 QY 1139 RPYDAIAGSILVDGKDLSKLNINSYRSLVYSQEPPTYOGTIRKENILGIVEDDVPEEF 1198
 DB 1068 RYHPLAGIVTMDGLDLSMDMSAIRNOYALVDPQETLFGQTIIRNLLGLDASKYSQBE 1127
 QY 1199 LIRACKDANIYDFINSLDEGFNTVYVSKGMLSGQOKORVALARALLRDPKILLDEATS 1258
 DB 1128 LEICKDANIIDFIRSLPLGFMTGCGKGNFSGQOKORVALARALLRDPVLLDEVTS 1187
 QY 1259 ALDSESEVQAALDAAARGRTTIVAHRLSTIQADYIYVDDQKIYESGHSLSVQOK 1318
 DB 1188 ALDSESQVQAALDQAAKQRTTIVAHRLSAVRNADICFLEDGVTIESGTHAELIRRR 1247
 QY 1319 GRYYELVNLQSLGK 1332
 DB 1248 GRFYAMLSQNIER 1261
 RESULT 6
 AAM62871
 ID AAM62871 standard; Protein; 1408 AA.
 XX
 AC AAM62871;
 XX
 DT 29-SEP-1998 (first entry)
 XX
 DE Multiple drug resistance protein called CneMDR1.
 XX
 KW Multiple drug resistance protein; CneMDR1; Identification:
 KW antifungal compound; resistance.
 XX
 XX Cryptococcus neoformans.
 PN
 XX US5786463-A.
 XX
 PD 28-JUL-1998.
 XX
 PF 08-MAR-1996; 96US-0612521.
 XX
 PR 08-MAR-1996; 96US-0612521.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Peery RB, Skatrud PL, Thornewell SJ,
 XX
 DR WPI: 1998-436610/37.
 DR N-PSDB: AAV42347; AAV42348.
 XX
 PT Dna encoding multiple drug resistance protein of Cryptococcus
 PT neoformans - also related vectors and transformed host cells, and
 PT method for expression of the protein
 XX
 PS Claim 1; Columns 17-26; 19pp; English.

D	b	143	KIRKPFHAIIMNOIGFVDHVGELNTRLTDVSKINIBGKICGMFQOMATPFGGI	202
Q	y	243	IAYVKWKLALICSSITVALVITWGGSGOFIYKSKSDSYGAGGVAEVIISIRNAT	302
D	b	203	IGFRGMKLTLLVILAIIPVLGSLGAIWAKILISFPTDKELQAYAKAGVAEVLAVYI	262
Q	y	303	AFGRQDLAQYEVHIDAEKWKGNQIYVGFHIGAMFGMLSNYGLGWMSRFLYDGA	362
D	b	263	AFGGQKLELEYNNINLEEARLIGIKATVAINISMGAFLILYASVLAWYGTSLVSK	322
Q	y	363	VDVODITLVMALILISFSISGIVSNPAAQFTNNAVAARJFGIDOSPDLVSMNGKL	422
D	b	323	YSIQOVLTFPAVVLIGFSGQASPNTEAFANRGAAYELFNIDDKPSIDSESKNYR	382
Q	y	423	DHEGHTLRNKAHYSREBYTMEVDVLSMBAGTKTALVGPSSGSKSTVGLYEFYM	482
D	b	383	DNIGNLEFKNIHPSYRSRDOYQILGNLKNYGSQTVALVGNSSGKSTVQLQRLD	442
Q	y	483	PVRGTVLDDHDKDLRLRQOISLVSEDPVLFEGTIIYKNIHGLISTKYNESEDRV	542
D	b	443	PTEBVASIDODIRTIWRLRLREIGVSEDPLEFATTAENIRYG----RENYTMB-	496
Q	y	543	RELLENAKANAMDFTTALPEGETNVOGRGLSGGQKORIALARAVSDPKILLDE	602
D	b	497	---IEKAVKEANADPEIMKLPHEFDYVBERGQSLSGGQRIALARAVNRPKILLDE	553
Q	y	603	ATSLADTKSEGVQOALERAEGRTTVIAHRLSTIKTANHYVLVNGKILAGQTHDELY	662
D	b	554	ATSLADTSESEVVOALDKAREGRTTVIAHRLSTIRNMDIAGFPGGVYVQGNHELM	613
Q	y	663	DRGAYIKLYBAQRINEQKADALDEDLTNADIANKITASSASSDLDKPTTIDRT	722
D	b	614	REKIYVETLWOTYAGNEIELGNVEGSKNEIDNLMMS---SKDSASSLIRRRSTRRSR	670
Q	y	723	GTHRSVSAILSKRPPTETKYSLMTLTKVVASFNPELIPYMLIGVYVLAGGGOPQA	782
D	b	671	GPHODKRLSTKEALDEBVDVPISFWRLK---LNSSEMPYVFGFCIVAGLQPAFS	726
Q	y	783	VLVKAISTSLPESOYSKLRHADPWSLMFFVVGIIQIETOSTNGAFAAVCSERLIRA	842
D	b	727	IIFSKVGVFTRNMTDETK-RHDSNLFSLFLILGIVISFTFFLGFTEGKAGELLTRL	785
Q	y	843	RSTAFRTLLQODIAFPKREKNSGALTSPFLSTETKHLGCVGYTLGTIIMSTTGAAT	902
D	b	786	RYWVFKSMLODQSWEDNPKNITGALTTRLANAGOVKATGARLAVITQIANJGTGII	845
Q	y	903	IATALIGKLLVYCSVYVVLACGFREYVLAQFGRSLAVEGSAFNCERTSSIRYA	962
D	b	846	ISLTYGQTLTLLAIYPTIATAGVYBEMKLSQALDKKELEGSKITATEIENFRYV	905
Q	y	963	SLTRERDWEIYHAQDLAQRSTLSIVLSSSLYASSQALVEFCVALGFWYCGTLLGHE	1022
D	b	906	SLTRERQEFEMWYQSLQIPRYNMLKKAHFGLITFSPTQAMNFYVLAACRFAYVABL	965
Q	y	1023	YDIFREYVCESEILFGAQSAGIYFSFAPDNGKAKNAAEFRRLPDKPOIDMWSEGEKL	1083
D	b	966	MTEFNVLVFSALYFGAMAAQVSSAFADYAKAKAYSASHIMIIEVPISIDSYSTGGLKP	1025
Q	y	1083	ETVGELEEFENHPRPTPREQVULVGLDITVYRQYVALVNPSCGCKSTTALLERFYD	1142
D	b	1026	NTELEGNKEVEYVNFTRDPIVDLQGLILEYKKGTTALVAGSSGGKSTVQLLREFYD	1085
Q	y	1143	AIAGSLIVDCKDLSKLININSYRSLFSLVSEPTLYGTAIKENILGLIYEDVDEEFLKA	1202
D	b	1086	PMAQTVLDEKYNQNLVQMLRAHLDIGVQEPILFPCSLAENIATGDNRSRVYSDQETIRA	1145
Q	y	1203	CKDANIYDTMSLPEGEFTNVSGKGLMSGQKORAVAIARALLRDPKILLDEATSAJD	1262
D	b	1146	AKENIHQIFTESTPDKNFRYGVGKGQSLGSGQKORIALARALVRPHILLDEATSAJD	1205
Q	y	1263	ESEVVOQALDAARGRTTVIAVHRLSTIOKADYIVYPOGKIVYESGHSSELVQKKGRY	1322
D	b	1206	ESEVVOEALDKAREGRTTVIAHRLSTIONMDIIVLWNGYVKEHGHTQDLQAGIYF	1265

QY	13323	EUYNIQSIGK	1332
:	:	:	:
Db	1266	SMVSVQAGAK	1275
RESULT 8			
ID	AAE18971		
XX	AAE18971 standard; Protein: 1280 AA.		
AC	AAE18971;		
XX			
DT	21-MAY-2002 (first entry)		
XX			
DE	Human P-gp mutant, MDR1-E634K/E636K (6-C1).		
XX			
KW	Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;		
RN	haematopoietic cell; cytotoxicity; cytostatic; P-gp; mutant; muteln.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 634	/note= "Wild type Glu substituted with Lys"	
FT	Misc-difference 636	/note= "Wild type Glu substituted with Lys"	
XX			
PN	WO200210205-A2.		
XX			
PD	07-FEB-2002.		
XX			
PP	01-AUG-2001; 2001WO-US24560.		
XX			
PR	01-AUG-2000; 2000US-222313P.		
XX			
PA	(UNII) UNIV ILLINOIS FOUND.		
PI	Ruth A, Robinson I;		
DR	WPI; 2002-206182/26.		
XX			
PS	Claim 42; Fig 3; 47pp; English.		
CC	The present invention relates to an isolated mutant human P-glycoprotein molecule. The invention is used as vaccine. An antibody specific for the mutant P-glycoprotein is useful for determining whether a human tumour sample comprises tumour cells expressing mutant human P-glycoprotein. A recombinant expression construct comprising nucleic acid encoding mutant P-glycoprotein is useful for administering an increased amount of a chemotherapeutic agent to an individual with cancer, by transducing ex vivo haematopoietic cells from the individual with the recombinant expression construct, reintroducing the transduced haematopoietic cells into the individual and administering an increased amount of a chemotherapeutic agent without consequent haematopoietic cytotoxicity. The mutant P-glycoprotein is useful for treating and diagnosing cancer and for screening compounds for the capacity to disrupt binding of cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux from cells expressing the mutant P-glycoprotein. The present sequence is human P-glycoprotein (P-gp) mutant, MDR1-E634K/E636K (6-C1) (multidrug resistance).		
CC			
XX			
SO	Sequence 1280 AA;		
Query Match	37.1%;	Score 2493;	DB 23; Length 1280;
Best Local Similarity	42.1%;	Pred. No. 1.8e-19;	
Matches 536; Conservative 228; Mismatches 448; Indels 60; Gaps 9;			

0Y	94	ETIVNVSFPGIMAYKMKMLIMVSTICJALAAASTFORILM-----	137
Db	28	KEKPPVSVSEMYSSMWLMDKLVMVGGTAAIIGHGGLPMLMVPGEIMTDFPANAENED	87
0Y	138	-----OISYDEP-----DELTFKNVLYFVYLIGCFEVPYVSVGFITYGEHATOKI	184
Db	88	LMSNITRSDNDNDFGFMNLEEDMTKRAYYSIGANGVLYAAVIOVSFMCJLAAGQIHKI	147
0Y	185	REYLESILKONIGYFDKLCIAGEVYTRITADINLIDOGISEKVGTLTALATVYAFITIA	244
Db	148	RKOFPHAIMRQELGMPFVDHVGELINTRLTDVASKINEGICDGKIMFQSGMAFFPGFTVG	207
0Y	245	YVKYWKALILCSSTVALVYLTMGGSQFIIKYSKSLDSYGAGCYTAEBVISINNAFAF	304
Db	208	FTFRMKTTLVLIALSPVLIGISAAMVAKMLISSFPDKELIYANAAGAABEVLAAIKPTIAF	267
0Y	305	GTOQKLAKOYEVHIDAEKKNQTKNOIYMGFMIGAMEGLMYSNYGLCFWNGSRFLVDGAVD	364
Db	268	GGQKKELEERYKNKMLEEAKRGIKKAITANISIAALLIYASALAFYGTTLVLSGEYS	327
0Y	365	VGDIILYVLMALILGSFSLIGNVSPNAQAFYTAANAARAKIFGTIDROSPLDYSNECKTIDH	424
Db	328	IGOVLYEVSFVLIGAFSGVQASPSIEAFANARAGAAIEFKIIDNKPSIDYSKSGHKPDN	387
0Y	425	FEGHIELRNKHIYPSRPETVYMEDVSLSPACKTIALYVPGSGSKSTYVGLVEFYAPV	484
Db	388	IKGNLEFNHVFSPSSKREYKILKGNLKVOSQOTVALVNGSCSGSTYVOLMQRIDYPT	447
0Y	485	RGVYLLGDHDKIDLKMLNQOISIVSOBEVLGCTIYKNIRGILGTYENESBDKRE	544
Db	448	EGMVSVDQDRTITVNFLEIKIIGVSOBEVYLATTIAENIRYG-----RENTVDE---	499
0Y	545	LIENAAKANAAHDTYTLPRGETITVNGQORFELISGGOKORIALARAVSDPRTILLDEAT	604
Db	500	-IEKAVRKANAYDPTMKLPKHPDPLVGERAQOLSGGOKORIALARLVANRPRTILLDEAT	558
0Y	605	SALDTKSEGVQAALEPABGRTTYIAHRLSTIKTAHNIVLVNGKIAEGOTHELDYDR	664
Db	559	SALDTSESAVVQVALLDKARKGRTTYIAHRLSTIVRNADVLAGDDDCVIEKENHBLMKE	618
0Y	665	GGAVRKILEQRIKNEOKREADLBDADBEDLTNDIAKIKITASASSDLDGCKPTTIDRGT	724
Db	619	KGIYFKLTYMTAGNEVE--LENA-----ADKSKSKIDALKKMSNDRSLIRKNST	668
0Y	725	HKSVSALLSKRP-----PETPKYSLTMLTKFASFERPEIPMLGLVFSYLAGG	777
Db	669	RRSVNGSQAOORKLSTKEALDESIPTSPFKRIK-----INTWMPYFVVGYVCALLINGL	724
0Y	778	QPTQAVLYAKAISTLSLPESOYSKLRHADFWSLMFVVGIIIOFTQSTNGAFAVCSER	837
Db	725	QPAFIIHPSKILIGVFTRIDBPETK-RQNSMLFELIFLALGIIISFTIFFLOGTFPGAGEI	783
0Y	838	LIRARSTAPRTILRODIAFPRKEENSTGALTSFLSTFKHLSGVSGVTLGLTMTSTTL	897
Db	784	LTKRLRIYFERSMLRODVSWPDOKNTGTLATLRLNDAAYKGAIGSLRAVITONIANL	843
0Y	898	GAALITAIALIMKIALVCSIVPVLACGAFYREYMAOFOFSRSKLIAYEBSANFAEATS	957
Db	844	GGIIILISITYMQLTLLILLAVPITAIAGVVEKKMLSGQALKRKRELBGAKIATPALEN	903
0Y	958	IRTVASLTREDDWEIYHAOLDAGRTSLISVLRSSILYASSQALVFCVALGFMYGCTL	1017
Db	904	FRTVVSILTQEQKPEHMAQSIQVRYRNSLKRKAHIFGTFESFTQAMMYFSYACQFRGAVL	963
0Y	1018	IGHNHYDIFRFVQCFSELIRGQAGAGVPSFAPDMKKAKMAAEFRRLDRKPQIDNMSE	1077
Db	964	VAHKLMSEDFVLVFSAVFEGAMAVGVSSFADYDAKAKISAHHIIMILEKPLIDYST	1023
0Y	1078	EGEKLEVEGELEFRNHYFRPTPEQVYRGLDILVYKPOYALVAGPSGSGSTYIALL	1133
Db	1024	BCJAMPNLTIEGNVTGEGVYFNTPTRPDLPVLOGISLEVKKQOTIALVSGSGCGKSTYVQLL	1088
0Y	1138	ERFYDALINGSLVDGKDISKILNINSYSPSLVSOEPTLYOGTIKENIILGIVEDVDPEE	1197

```

Db 1084 ERFYPLAKVLLDGEIKRLNVQMLRAHGLVTSQEPILFDCSIAINAYGNSRVSOE 1143
OY 1198 FIKACKDANIYDFIMSLPEGFNTVVGSKGSLGOKORAVIARALNDPKLLIDEAT 1257
Db 1144 EIVRAKENIAFIIESLPNKSTKVGDKGTQSLGOKRIARALVQPHILLDEAT 1203
OY 1238 SALDSESEKVVQAALDAARGRTTIAVAHRLSTIQKADVIYVFDGKIVESGTHSELVOK 1317
Db 1204 SALDSESEKVVQDALDKAREGRTCIIVAHRLSTIQADLIIVFONGRVKHEGTHQOOLAO 1263
OY 1318 KGRYYELVLOS 1329
Db 1264 KGIYFSMVSQA 1275

RESULT 9
AAM44073
ID AAM44073 standard: protein; 1280 AA.
AC AAM44073;
XX
XX 26-JUN-1998 (first entry)
DT
XX Human multidrug resistance P-glycoprotein MDR1.
DE Human multidrug resistance P-glycoprotein; MDR1; prokaryotic homologue.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX FH 1.640
XX FT Region /note="MDR1-N from Fig 1"
XX FT 641..1280
XX FT Region /note="MDR1-C from Fig 1"
XX PM WO9740160-A1.
XX PD 30-OCT-1997.
XX XX 24-APR-1997; 97WO-N000216.
XX PF 24-APR-1996; 96EP-0201094.
XX XX 24-APR-1996; 96EP-0201094.
XX PA (UYGR-) RIJCSUNIV GRONINGEN.
XX PI Bolhuis H, Konings WN, Van Veen HW, Venema K;
XX WPI: 1997-535844/49.
XX PT Prokaryotic homologue of human multiple drug resistance protein -
XX used to screen for compounds that inhibit, or avoid, drug resistance
XX
XX Claim 10; Fig 1; 35pp; English.
XX
XX The present invention describes a recombinant or isolated nucleic acid
XX (1), derived from a prokaryotic gene, which encodes at least a specific
XX and/or functional part of a transporter protein (7P), or its
XX derivatives, which has functional and/or structural similarity with the
XX P-glycoprotein (PG) encoded by the human multidrug resistance
XX (MDR1) gene. The present sequence represents the human MDR1 protein,
XX derived from MDR1-N and MDR1-C as shown in the specification in
XX figure 1. (1) is used to express recombinant proteins; its fragments
XX are also useful as probes and primers for detection and amplification
XX of related DNA. The protein produced, or cells expressing them, are
XX used to determine if substances can inhibit, or avoid, MDR proteins,
XX and in a screening method for identifying compounds that inhibit
XX transport of cytotoxic substances from cells. Also, cells with a
XX transmembrane protein, especially where expressed from (1), can provide
XX (additional) MDR, particularly for use as a model system to study
XX mechanisms of action of PG.
XX
XX Sequence 1280 AA:

```

```

Query Match 37.0%; Score 2489; DB 18; Length 1280;
Best Local Similarity 42.1%; Pred. No. 3.8e-197;
Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

OY 94 EELKVNISFFGLMRYATKADILIMVISTICALMASTFORIMLY----- 137
Db 28 KEKKPTVSVFSMEFRYSNNKLYMVGTLAALIHGALPLMLVFEGMTDIFANGLNLED 87
OY 138 -----QISDEFEY-----DELTKNVLYFYVLGICEFVTVYVSGVITYGEHATOKI 184
Db 88 LMSNTNRSDINDTGFEMNLBEDMTNRYATYSGIGAVLVAAITOVSFCLAGROIHTI 147
OY 185 REYYLESILRONIGYEDKLGAGEVTRITADNTLIDQGISERYGLTLLATFVTAIA 244
Db 148 RKQFFAIRMROELGEMFDVADGELNRLDDVSKINEGIDKMGFMQSMATFFCFIYG 207
OY 245 YKRYKFLALICSTYVALVLMGGSQFLIKYSKSLSDSYGAGGVAAEVVISRNATAF 304
Db 208 FTRGMKLTIVLIAISPVLGLSAVWMAKILSFTDKELLAYAKAGVAEEVLAIKRTVIAF 267
OY 305 GTQDKLAKOYEYHIDEAEKMGTKNOIVMGFMIGAMGGLVSNVYGIFPMGSRFINDGAND 364
Db 268 GGOKKELEPNKKNLEBAKRIKIKATTANISIGAFLLIYASTALAFWGTLLVSGEYS 327
OY 365 VGDILTVMALLIGSESLGNVSPNAQFTNAAVAAAKIFGTIDROSPLDYPNECKTLDH 424
Db 328 IGOVLVFPFSVLIGAFVSGASPSIEAFANARGAAEIKRIIDNKRPSIDYSKSGKRPD 387
OY 425 FECHIELRNVKIHYSPRPVYTWEDVSLSPRAGKTTALVSGSGSKSVVGLVERPMV 484
Db 388 IKGNELEFRNVHSYSPRKVKIKLGLNLKVOSGQFALVGNSGCGKSTVQLMORLYDPT 447
OY 485 RCTVLLDGDIDOLNRLRQOISLVSOEPVFGFTTYINIRHGLTGTVENESSEDKVNE 544
Db 448 EGMVSVDSGDITRTINRFIRELIGVYSGEPVLEFATTIAENIRYG-----RENTYDDE--- 499
OY 545 LIENAKMANAHDFITALPEGEYETNVGREGFLLSGGOKRIAIARAVSDPKTILLDEAT 604
Db 500 -LEKAVKEANAVDFIMKLPHKFDTLVGERGAQLSGGOKRIAIARALVANPRTILLDEAT 558
OY 605 SALDSESEKVVQAALERAEGRTTIVIAHRLSTIKAHNIYVLYNGKIAEGSTHDELYR 664
Db 559 SALDSESEKVVQALDKAKGRTTIVIAHRLSTVRADVIAGFDGVITYEKENHDELMKE 618
OY 665 GGAYRKLVEAQRINQEKADALEDAEDLTNADIAKIKITASSASSADLDGKFTTIDRTCT 724
Db 619 KGIYFKLVIMQTAGNEVE---LENA-----ADESKSIIDALEKSSNDSRSILRKST 668
OY 725 HKVSASAILSKRP-----PETPKYSLMTLLKLFVASENREIPYMLIGLVFSYLAGGG 777
Db 669 RRSVRSQAQDRKLSYKELADESIPVSPWRIMK-----LNTWEMPVYGVGCATINGL 724
OY 778 OPTQAVLYAKAISTLSLPSQYSKLRHDDFWMSLFWFVVGIIQFTIQSTNGAFAVCSER 837
Db 725 QPFAFATFSKIIIGVFRRIDDPETK-RQNSNLFSLLFLAGIISFTFFLGQTFFGKAGI 783
OY 838 LIRRASTFPTILRODIAFPDKEENSGALPSFLSTERRKHSVGSVGLGITLMTSTL 897
Db 784 LKRLRLRYMFRBMLKQDVSWFDDPKRTTALTRRLANDAAQKGAIGSLANTYITONIANL 843
OY 898 GAIIITIALIGKMLALVCISVVPVLLACGYFRPYMLAQFOSRSKTLAYEGSANFACEATSS 957
Db 844 GGIITISFYGMQLFTLLAIVPIIAIAGVEMKMLSCGALKDKELBEGAGIATEAIEN 903
OY 958 ITFVASLTERVWYELIYNAQDLAOGRTSLISYRSLSLYASSQALVFCVALGMYGTL 1017
Db 904 FRTVYSLTQEQFEHRTAQSLOVPRYNSLRKAHIGITFISFOQAMMYFSYACCFRGAFL 963
OY 1018 LGHHEYDIDRFVCFSEILFGAOSAGTVSFAPDMGKANAAEFRRRLDPRQDIDNNS 1077
Db 964 VAHKLSMFEVDVLVFSAVVFGMAAVGOVSSFADYAKAKAISAHITIMITEKTPLLIDYST 1023

```



```

Db      844 GTGIIISFIYQMQLTLLLAIVPIALAGVEMKMSGALKDKKELEGAGKITEAIEN 903
QY      958 IRTVAISIRRDQWEIYTHAOLDAGRTSLISVLRSSLLYASSQALVFCVAGWYIGTL 1017
Db      904 FRTVSLTQOKKEEHMYAQSLOVPPYRNSLKAHIFGITFSFTQAMVFSYAGCFRGAYL 963
QY      1018 LGHHEVDIFFEFCSEILGASAGTGFSPADPMGAKANAAEFRLFPKRPQIDMSE 1077
Db      964 VAKKMSFEVVLVFSNVGAMAVGVSSFADPYAKAKISAAHIIITEKPLIDSYST 1023
QY      1078 EGEKLETEGEIEEFRRVHYRPTPRPQVLRGLDLTVKPGQYVALVGPSCGKSTTALL 1137
Db      1024 EGLMPPTLEGNVTEGEVFNYPTRPDIPVLOGLSLVKKGQTLALVSSGCGKSTVQOLL 1083
QY      1138 ERTYDAAGSLVDGDKSLNINSYSEFLSYSQEPTLQGTIKENIILGIYDDVPEE 1197
Db      1084 ERTYDPLAGVLLDGRKIKLNQWMLAHGLIVSQEPILDFCSIAENIAVGDNSRVVSOE 1143
QY      1198 FLIKACKDANIYDFIMSLPGFNTVVGSKGMLSGGOKORVAIARALLRDPKILLDEAT 1257
Db      1144 EYRAAKENAIHAFIESLPKRYSTKVGDGKTOLSGGOKHAIARALVROPHILLDEAT 1203
QY      1258 SALDSEKVVQAALDAARGRTTIAVAHRLSTIQKADVIYVDOGKIVESGTHSELVOK 1317
Db      1204 SALDTESEKVVQEBALDKARGRCTIVIAHRLSTIQNADLIYVFOGKHVKEGHTHQLLAQ 1263
QY      1318 KGRYVELVNGS 1329
Db      1264 KGIYFSMVSVQA 1275

```

RESULT 12

AAEI8964 standard; Protein; 1280 AA.

ID AAEI8964:

AC AAEI8964:

DT 21-MAY-2002 (first entry)

DE Human wild type P-glycoprotein (wt P-gp).

KW Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer; haematopoietic cell; cytotoxicity; cytostatic; P-gp.

OS Homo sapiens.

PN WO200210205-A2.

PD 07-FEB-2002.

PE 01-AUG-2001; 2001WO-US24560.

PR 01-AUG-2000; 2000US-222313P.

PA (UNII) UNIV ILLINOIS FOUND.

PI Ruth A, Roninson I;

XX WPI; 2002-206182/26.

XX New human P-glycoprotein mutants for treating cancer, comprises an

PT ability to confer increased resistance to chemotherapeutic drugs

PT relative to wild type P-glycoprotein or P-glycoprotein having a glycine

PT to valine substitution

PS Claim 44; Fig 3; 47pp; English.

CC The present invention relates to an isolated mutant human P-glycoprotein

CC molecule. The invention is used as vaccine. An antibody specific for the

CC mutant P-glycoprotein is useful for determining whether a human tumour

CC sample comprises tumour cells expressing mutant human P-glycoprotein. A

CC recombinant expression construct comprising nucleic acid encoding mutant

P-glycoprotein is useful for administering an increased amount of a chemotherapeutic agent to an individual with cancer, by transducing ex vivo haematopoietic cells from the individual with the recombinant expression construct, reintroducing the transduced haematopoietic cells into the individual and administering an increased amount of a chemotherapeutic agent without consequent haematopoietic cytotoxicity. The mutant P-glycoprotein is useful for treating and diagnosing cancer and for screening compounds for the capacity to disrupt binding of cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux from cells expressing the mutant P-glycoprotein. The present sequence is human wild type P-glycoprotein (wt P-gp).

Sequence 1280 AA:

Query Match 37.0%; Score 2489; DB 23; Length 1280;

Best Local Similarity 42.1%; Pred. No. 3.8e-197; Mismatches 448; Indels 60; Gaps 9;

Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

```

QY      94 EEIKVNISFEGLMRYATKMDILINVIITCAIAASTFORIMLY----- 137
Db      28 KKKPTVSVFEMFRYSNMLDKLYMVGTLAIHIGAGLPLMLVGEWTFDIFANAGNED 87
QY      138 -----QISDEFT---DELTKNLYFVYIGIEFTVYVSTVGFIYTGHAPOKI 184
Db      88 LMSNITNRSNDINDGFENLDEDMTRYVYVYSGIAGVLAAYIOVSFMCIAAGROIHKI 147
QY      185 REYLESILRONIGYFPKLGEVTTTRTADTNLIDGSIKVGILTLATFTYATFIA 244
Db      148 RKQFFHAIKROELGFWPDHVDGELNTRLDVSKINSIGIKIGFQSMATFTFTGTVG 207
QY      245 YKRYWKALICOSTIVALVLTMGSGGQFIIRYKSKSDSYGAGTVAEYVSSIRNATAF 304
Db      208 FTRGMKTLIVLAIISPVLGSLAAVMAKILSFTEPKELLAYAKAGVAEVLAIARTVIAF 267
QY      305 GTQOKLAKQYEVHLDPAKMGCTKNOIYMGEMTIGAMFGLMYNTYIGFWMGSRFLVDGAVD 364
Db      268 GQQRKELEERYKNNEEAKRIGIKRAITANISIGAFLLIYASALAFYVGTLLVSGEYS 327
QY      365 VGDILTVMALILIGSFSIGVSNPNAQFTNAVAAAKIFGTIDROSPLDPPYNSNEKTLDH 424
Db      328 IGVITVYFSSVLIGAFSFGQASPSIEAFANARGAAYELFKIITDNKPSIDSTSKSGHKPDN 387
QY      425 FECHIELNRVNHIIYPSREVTVMEDVLSMPAGKTTALVPSGSGKSTVGLIVERFYMPV 484
Db      388 IKGNLEFRNVHFSYPSRREVKIKGLNLKYGSGGVVALVNGSGGKSTVQLMQGLYPT 447
QY      485 RGYVLDSGDKLNDLRLRQOISLYSQEPVLEGTITKYNRHGLIGIKYENESDQVRE 544
Db      448 EGMVSVDDQDIRTINVRLEIIGVSOEPLVFTTIAENIRYG-----RENVYVDE--- 499
QY      545 LIENNAKMANAHDFTLPEGETNVGGRFLSGGOKORIAARAVVSDPKILLDEAT 604
Db      500 -IERAVKEMANVDFIMKLPKFDVLVGERGQOLSGGOKORIAARALVRNPKILLDEAT 558
QY      605 SALDTKSEGVQAALERAEGRTTIVIAHRLSTIKTANIVLVNGKIAEOGTHDELVDYR 664
Db      559 SALDTESEAVVQVLDKARKGRRTTIVIAHRLSTVYANADVIAGFDGVLVEKGNDELKKE 618
QY      665 GGATRKLYEAGRINEQKADALDADADLTNADAKIKITASSASDDGKPTTIDRKGT 724
Db      619 KGIYFKLVMTQAGNEV---LENA-----ADESKSEIDALENSMDSSSLRKST 668
QY      725 HKVSSATLSLRP-----PETPKYSLWTLTKVASFENREIYMYLIGVFSVLAGG 777
Db      669 RRSVAGSOQDRKSLTKALDESDIPVYFWRMK-----LNLTEPFIYVVGFCALINGL 724
QY      778 QPTQAVLYAKAISTLSLPEISOYSKLRHADFWSLMFVVGIIQFTTQSTNGAAPVCSER 837
Db      725 QPAFAIIRSKILIGVTRIDDPETK-RQNSNLSFLIFLAGIISFTFTFLOGTFPKAGEI 783
QY      838 LIRARSTAFITILRODAFADKENSFGALTSFTSTKHSVGSVGTLLGTLMTSTL 897
Db      784 LTKRLRYVFRSMILKQDVSWFDDPKNTTGALTTRLANDAAOVKAIGSLAVITONIANL 843

```



```
Db 619 KGIFYELVTMOAGNFKVK---LENA-----ADKSKSKIDALKMSSNDRSSILIRKST 668
OY 725 HKSVSAILSKRP-----PETPKYSLMTLLKFVASFNRPEIPYMLIGLVSVLACGG 777
Db 669 RRSVRSQOQDRKLSTKEALDESIPVYSFWRIK----LNLTEPYPYVGVFCAILINGCL 724
OY 778 OPTQAVLVAKAISTLSLPSQYSKLRHDADFWSLMEFFVGGIIQFTQSTNGAFAVCSER 837
Db 725 OPAFAIIFSKIIIGFTRIDDEPK-RONSMLFSLFLALGIIISFTTEFLGFTFGKAGEI 783
OY 838 LIRRASTAPRTLLRODIAFPDKEENSTGALTJFSLSTETKHLGSGCVTLGTLMTSTYL 897
Db 784 LKRLRYMWFERSMLRODVSFWFDDPKNTGTGALTTRLANDAAQVKAIGSLAVITONIANL 843
OY 898 GAIIITAIIGMTALVCISVVPVLLACGFYRPMLAQFQSRSKLAVESGANFACEATSS 957
Db 844 GGIITISFYGMOLJLLLAIVPIIAIAGVVEKMKLSGALKDKKELEGAKIATEAIEN 903
OY 958 IRTVASLTRERDVWEIYHAQLDAQRTSLISVLRSSSLVASSQALVEFCVALGFYGGTL 1017
Db 904 FRTVVSILTOEQKEHMYAOSLOVPYRNSLKAHIFGITFSFTQAMMYFSYAGCFREGAYL 963
OY 1018 LGHHEVDIRFEVCESEILFQAQSACTVFSFADMGKAKANAAEFRRLEFDKPOIDNWSE 1077
Db 964 VAKHIMSFEEDVLLVFSAVVFGAMAVGOVSFADPVAKAKISAHHIIMIIETPLIDSYST 1023
OY 1078 EGEKLETVGEIEFRNVHFPTRPEOPVLRGDLTVKPGQYVALVPGSGCGKSTTIAL 1137
Db 1024 EGIIMPNTLEGNTVEGEVFNYPTRPDIPVLOGLSLEVKKGOTLALVGSSCGSKSTVQTL 1083
OY 1138 EREYDAIAGSIIVDGDISKLNINSYRSLVLSVSOEPTLYOGTIKENILGIYEDDVPEE 1197
Db 1084 EREYDPLAGVLLDGEIKRILNOMLRAHLGIYSOEPILEDCSIAENIAYGDNSRYVSOE 1143
OY 1198 FLIKACKDANIYPIFMSLPEGENTVYSGSGMLSGGOKORVATARALLRDPKILLDEAT 1257
Db 1144 EIVRAKENANIHAFIESLPKYSTKVGDKGTOLSGGOKORIAIARALVRQPHILLDEAT 1203
OY 1258 SALDSESEKVVQALDPAARGTTIAVHRSLSTIOKADVIYVFDGKIVESGTHSELVOK 1317
Db 1204 SALDTESEKVVQALDKARBGRTCIYIAHRLSTIIONADLIVFONGRVRKEHGHQOLLAQ 1263
OY 1318 KGRYELVNLQS 1329
Db 1264 KGIFYFSMVSVOA 1275
```

Search completed: April 1, 2003, 16:00:39
Job time : 91 secs

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that this is essential for ensuring transparency and accountability in the organization's operations.

2. The second part of the document outlines the various methods and tools used to collect and analyze data. It highlights the need for a systematic approach to data collection and the importance of using reliable sources of information.

3. The third part of the document describes the process of interpreting the data and drawing conclusions from it. It stresses the importance of considering all relevant factors and avoiding biases in the analysis.

4. The fourth part of the document discusses the implications of the findings and the steps that should be taken to address any issues identified. It emphasizes the need for a proactive approach to problem-solving and the importance of continuous improvement.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:59:08 ; Search time 20 Seconds
(without alignments)
1962.509 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLETNPLSPETAMREPAAE.....YOKGRYELVNIQSLRGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6721	100.0	1334	2	US-08-996-545-2
2	6721	100.0	1334	4	US-09-328-320-2
3	5407.5	80.5	1349	2	US-08-612-734B-2
4	3992	59.4	1307	1	US-08-395-246C-2
5	2682.5	39.9	1408	1	US-08-612-521-2
6	2478	36.9	1280	6	5206352-4
7	2477	36.9	1280	2	US-08-583-276-19
8	2469.5	36.7	1279	2	US-08-784-649A-2
9	2466	36.7	1280	2	US-08-752-447-2
10	2466	36.7	1280	4	US-09-316-167-2
11	2408.5	35.8	1275	4	US-09-120-513-2
12	2408.5	35.8	1275	3	US-09-450-105-2
13	2208.5	32.9	1308	2	US-08-996-644-2
14	2208.5	32.9	1308	3	US-09-352-552-2
15	1991.5	29.6	1302	1	US-08-232-537-2
16	1940	28.9	1263	4	US-09-351-224E-11
17	757	11.3	791	1	US-08-394-880B-2
18	746.5	11.1	1622	4	US-08-972-927-6
19	739.5	11.0	1261	4	US-09-605-785-538
20	739.5	11.0	1261	4	US-09-439-313-538
21	738	11.0	1621	4	US-08-972-927-3
22	711	10.6	1528	1	US-08-463-109A-6
23	711	10.6	1528	2	US-08-462-109B-6
24	711	10.6	1528	2	US-08-460-907B-6
25	711	10.6	1528	3	US-08-463-179A-6
26	711	10.6	1528	3	US-08-461-384B-6
27	701.5	10.4	580	4	US-09-134-001C-5611

28	701	10.4	327	1	US-08-463-092B-9	Sequence 9, App11
29	701	10.4	327	2	US-08-460-907B-9	Sequence 9, App11
30	677.5	10.1	1531	1	US-08-463-092B-4	Sequence 4, App11
31	677.5	10.1	1531	2	US-08-462-109A-4	Sequence 4, App11
32	677.5	10.1	1531	2	US-08-460-907B-4	Sequence 4, App11
33	677.5	10.1	1531	3	US-08-463-179A-4	Sequence 4, App11
34	677.5	10.1	1531	3	US-08-461-384B-4	Sequence 4, App11
35	676	10.1	748	4	US-09-061-764A-19	Sequence 19, App1
36	674.5	10.0	1437	3	US-09-061-400-2	Sequence 2, App11
37	674.5	10.0	1453	3	US-09-001-273-2	Sequence 2, App11
38	674.5	10.0	1453	4	US-08-843-459A-2	Sequence 2, App11
39	670	10.0	1228	4	US-09-605-785-537	Sequence 537, App
40	670	10.0	1228	4	US-09-439-313-537	Sequence 537, App
41	669.5	10.0	1531	1	US-08-141-893-2	Sequence 2, App11
42	669.5	10.0	1531	1	US-08-463-092B-2	Sequence 2, App11
43	669.5	10.0	1531	2	US-08-462-109B-2	Sequence 2, App11
44	669.5	10.0	1531	2	US-08-460-907B-2	Sequence 2, App11
45	669.5	10.0	1531	3	US-08-463-179A-2	Sequence 2, App11

ALIGNMENTS

```
RESULT 1
US-08-996-545-2
; Sequence 2, Application US/08996545
; Patent No. 5928898
;
GENERAL INFORMATION:
;
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
;
TITLE OF INVENTION: Multiple Drug Resistance Gene atpd of
;
NUMBER OF SEQUENCES: 3
;
CORRESPONDENCE ADDRESS:
;
ADDRESSEE: Eli Lilly and Company
;
STREET: Lilly Corporate Center
;
CITY: Indianapolis
;
STATE: Indiana
;
COUNTRY: U.S.
;
ZIP: 46285
;
COMPUTER READABLE FORM:
;
MEDIUM TYPE: Floppy disk
;
COMPUTER: IBM PC compatible
;
OPERATING SYSTEM: PC-DOS/MS-DOS
;
SOFTWARE: PatentIn Release #1.0, Version #1.30
;
CURRENT APPLICATION DATA:
;
APPLICATION NUMBER: US/08/996,545
;
FILING DATE:
;
CLASSIFICATION: 435
;
ATTORNEY/AGENT INFORMATION:
;
NAME: Webster, Thomas D.
;
REGISTRATION NUMBER: 39,872
;
REFERENCE/DOCKET NUMBER: X-11766
;
TELECOMMUNICATION INFORMATION:
;
TELEPHONE: 317-276-3334
;
TELEFAX: 317-276-2763
;
INFORMATION FOR SEQ ID NO: 2:
;
SEQUENCE CHARACTERISTICS:
;
LENGTH: 1334 amino acids
;
TYPE: amino acid
;
TOPOLOGY: linear
;
MOLECULE TYPE: protein
;
US-08-996-545-2

Query Match 100.0%; Score 6721; DB 2; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPLETNPLSPETAMREPAAETSTEEQASTPHAADEKKIISDLSAPSTATTAPDKRRP 60
Db 1 MSPLETNPLSPETAMREPAAETSTEEQASTPHAADEKKIISDLSAPSTATTAPDKRRP 60
```

Qy	61	KSSSSNNAVSNEVDALIAHLPEDEROVLTQTOLEIKVINSFGLMRVATKMDILTMVTS	120
Dp	61	KSSSSNNAVSNEVDALIAHLPEDEROVLTQTOLEIKVINSFGLMRVATKMDILTMVTS	120
Qy	121	TICAIAAASTFORIMLYOISDEYFDELTKNVLIFYVLGIGEFVTVVSVTFGIYGEHA	180
Dp	121	TICAIAAASTFORIMLYOISDEYFDELTKNVLIFYVLGIGEFVTVVSVTFGIYGEHA	180
Qy	181	TOKIREYVESLILNQNGYEDKDGAGVVTTRITADINLJODGISEKVGLTITRAIFVYA	240
Dp	181	TOKIREYVESLILNQNGYEDKDGAGVVTTRITADINLJODGISEKVGLTITRAIFVYA	240
Qy	241	FIIAVYKWKMLALICSSITVYALVTMGGSQOIFIKKSKLSDSYGAGTVAEEVISIRN	300
Dp	241	FIIAVYKWKMLALICSSITVYALVTMGGSQOIFIKKSKLSDSYGAGTVAEEVISIRN	300
Qy	301	ATAAGTODKLAKOYEVLHDEAEKKGITNOIYMGFMIGAFGLMTSNYGLGFMGMSFFLD	360
Dp	301	ATAAGTODKLAKOYEVLHDEAEKKGITNOIYMGFMIGAFGLMTSNYGLGFMGMSFFLD	360
Qy	361	GAYVGDILVYMAIILGSGFSLGVSPNAQAFINAAVAALKITGTIDROSPDOPYNEEK	420
Dp	361	GAYVGDILVYMAIILGSGFSLGVSPNAQAFINAAVAALKITGTIDROSPDOPYNEEK	420
Qy	421	TLDHFEGBIELRNKHIYPSRPEVTYVEDVSLSPACKTTALVPSGSGKSTVVGLYVER	480
Dp	421	TLDHFEGBIELRNKHIYPSRPEVTYVEDVSLSPACKTTALVPSGSGKSTVVGLYVER	480
Qy	481	YMPRGVYLDGHDIKOLNLRMLROOISLVSOSPVLFGTTIYNIRHGLIGTFYEMESBD	540
Dp	481	YMPRGVYLDGHDIKOLNLRMLROOISLVSOSPVLFGTTIYNIRHGLIGTFYEMESBD	540
Qy	541	KVRELLENAAKMANAHDFITALPEGYETNYGORGFLISGGOKORIAIARAVVSDPKITLL	600
Dp	541	KVRELLENAAKMANAHDFITALPEGYETNYGORGFLISGGOKORIAIARAVVSDPKITLL	600
Qy	601	DEARSAIDTSGSEGVVQALERRAEGRTTYIARLSTIKTAHIYVLVNGKIAEOGTBHE	660
Dp	601	DEARSAIDTSGSEGVVQALERRAEGRTTYIARLSTIKTAHIYVLVNGKIAEOGTBHE	660
Qy	661	LVDRGAYRKLVENQRINEQREADALEDDADAEJLTNADIAKIKTASASASDDJGKPTTID	720
Dp	661	LVDRGAYRKLVENQRINEQREADALEDDADAEJLTNADIAKIKTASASASDDJGKPTTID	720
Qy	721	RTGTHKSVSSAILSKRPETTPKYSILMTLTKFYASFRPRPITPMLGLYFVSVLAGGQPT	780
Dp	721	RTGTHKSVSSAILSKRPETTPKYSILMTLTKFYASFRPRPITPMLGLYFVSVLAGGQPT	780
Qy	781	QAVLYAKAISLSPESQYSKLRHADPWSLMEFVVGIIOTQSTNGGAFAVCSRLIR	840
Dp	781	QAVLYAKAISLSPESQYSKLRHADPWSLMEFVVGIIOTQSTNGGAFAVCSRLIR	840
Qy	841	RARSTAFRTILRODIAEFDEENSTGALTJFSLTEKHLISGVSGLTILMTSTTIGAA	900
Dp	841	RARSTAFRTILRODIAEFDEENSTGALTJFSLTEKHLISGVSGLTILMTSTTIGAA	900
Qy	901	IIITLALIGMKLALVCISVVPVLLACGFRYRMYMAQOSRSKLAYESANFACATYSITF	960
Dp	901	IIITLALIGMKLALVCISVVPVLLACGFRYRMYMAQOSRSKLAYESANFACATYSITF	960
Qy	961	VASLTRENDWEIYHAOLDQAGRTSLTSVLRSSILVASSQALFFCVALGFVWGGLLGH	1020
Dp	961	VASLTRENDWEIYHAOLDQAGRTSLTSVLRSSILVASSQALFFCVALGFVWGGLLGH	1020
Qy	1021	HEYDIFRFVCFSSITLGAOSAGITVSPADMGKAKANAAEFRRJLDRKQIDNMSSEGE	1080
Dp	1021	HEYDIFRFVCFSSITLGAOSAGITVSPADMGKAKANAAEFRRJLDRKQIDNMSSEGE	1080
Qy	1081	KLEIVFEGIEBRNHFYPTRPBPQVYLRGLDILVYKGOVYALVGPGGCKSTTIALLEFR	1144
Dp	1081	KLEIVFEGIEBRNHFYPTRPBPQVYLRGLDILVYKGOVYALVGPGGCKSTTIALLEFR	1144

QY	1141	YDIAAGSIIVGDGDKIKLNINSTRPSLTVSOSOPPTYOCTIKENTILLIGVEDDVPEEFLI	1200
Dd	1141	YDAIAGSIIVGDGDKIKLNINSTRPSLTVSOSOPPTYOCTIKENTILLIGVEDDVPEEFLI	1200
QY	1201	KACKDANIYDFIMSLPEGFNTVVGSKGMI.SGGOKORVAIAFALLRDPKILLILDEATSAL	1260
Dd	1201	KACKDANIYDFIMSLPEGFNTVVGSKGM.LSGOGORVAIAFALLRDPKILLILDEATSAL	1260
QY	1261	DSEBEKVQAALDAAARGRTTAVAHRLSTIQKADVIYVFDDGKTIVESGTISELYÖKKGR	1320
Dd	1261	DSEBEKVQAALDAAARGRTTAVAHRLSTIQKADVIYVFDDGKTIVESGTISELYÖKKGR	1320
QY	1321	YYELVYNLOSIGRGH	1334
Dd	1321	YYELVYNLOSIGRGH	1334

```

RESULT 2
US-09-328-320-2
: Sequence 2, Application US/09328320
: Patent No. 6228615
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Waard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/328,320
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/996,545
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11766
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1334 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-328-320-2

Query Match 100.0%; Score 6721; DB 4; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPETNPSPETLMREPASTSTTEEQASTPHAADEKILDSLAPSSSTTATPADKEHRP 60
1 MSPETNPSPETLMREPASTSTTEEQASTPHAADEKILDSLAPSSSTTATPADKEHRP 60
1 KSSSSNNAVSNEVDALIAHPEDEROYLKTQLEETIKYNISFFGIMWRATKDIILNIVIS 120

```

```

Db      61 KSSSSNNAVNEVDALLAHLPEDEROVLKTOLKEIKVNIISFGIMRIVTKMDILIMVIS 120
Qy      121 TICAIAAATFORIMLYOISYDEFDYELTKNLYEYLLIGEFVYVYVSGEITYGEHA 180
Db      121 TICAIAAATFORIMLYOISYDEFDYELTKNLYEYLLIGEFVYVYVSGEITYGEHA 180
Qy      181 TOKIREYLESILRONIGYFDKLGAEVTTTITADTNLIQDGISEKVGTLTALATFVTA 240
Db      181 TOKIREYLESILRONIGYFDKLGAEVTTTITADTNLIQDGISEKVGTLTALATFVTA 240
Qy      241 FIATAVYKWKALLICSSITVALVLMGGSSOFTIKSKSDSYGAGGVAAEVISSIRN 300
Db      241 FIATAVYKWKALLICSSITVALVLMGGSSOFTIKSKSDSYGAGGVAAEVISSIRN 300
Qy      301 AFAFGTODKLAKOYEYHDEAEKMGTKNOIYMGFMIGAMGLMYNGYGFMMWGSRELV 360
Db      301 AFAFGTODKLAKOYEYHDEAEKMGTKNOIYMGFMIGAMGLMYNGYGFMMWGSRELV 360
Qy      361 GAVDVGDIITVLMALILIGSFSLGNSVPMNAQAFITNAVAAAATFCTIDROSPLDPYSNEGK 420
Db      361 GAVDVGDIITVLMALILIGSFSLGNSVPMNAQAFITNAVAAAATFCTIDROSPLDPYSNEGK 420
Qy      421 TIDHEBGHIELKNVKHITSPREVTYMEDVSLSPAGKTTALVGPBGSGKSTVGLVERF 480
Db      421 TIDHEBGHIELKNVKHITSPREVTYMEDVSLSPAGKTTALVGPBGSGKSTVGLVERF 480
Qy      481 YMPVKGTVLLDGHDKDLMLRMLROOISLVSOEPVLFGTITKINRHLIGTYEENESD 540
Db      481 YMPVKGTVLLDGHDKDLMLRMLROOISLVSOEPVLFGTITKINRHLIGTYEENESD 540
Qy      541 KYRELIENAKAMAHDFITLALPEGETVNGOFGFLISGOKORLAIARAVSDPKILL 600
Db      541 KYRELIENAKAMAHDFITLALPEGETVNGOFGFLISGOKORLAIARAVSDPKILL 600
Qy      601 DATSALDTPKSGVVOALERAAGERTTIVIAHRLSTITAHNIVLVNGKIAEOGTHDE 660
Db      601 DATSALDTPKSGVVOALERAAGERTTIVIAHRLSTITAHNIVLVNGKIAEOGTHDE 660
Qy      661 LVDRGAGYKRLVEAORINOEKADAELEDAEDLTNADIAKITASASSDLDGKPTTD 720
Db      661 LVDRGAGYKRLVEAORINOEKADAELEDAEDLTNADIAKITASASSDLDGKPTTD 720
Qy      721 RGTGHSVSSAILSKRPETTPRYSLMTLLKFVASFNRPEIPYMLIGLVFSVLAGGQPT 780
Db      721 RGTGHSVSSAILSKRPETTPRYSLMTLLKFVASFNRPEIPYMLIGLVFSVLAGGQPT 780
Qy      781 QAVLAKAISTSLPESQYSKLRHADFWSLMEFVVGIIQFTIQTSTNGAFAVCSERLIR 840
Db      781 QAVLAKAISTSLPESQYSKLRHADFWSLMEFVVGIIQFTIQTSTNGAFAVCSERLIR 840
Qy      841 RARSTAFRTILRODIAFPFKENSTGALSTETKHLISGVSGVTLGTITMSTLGA 900
Db      841 RARSTAFRTILRODIAFPFKENSTGALSTETKHLISGVSGVTLGTITMSTLGA 900
Qy      901 IITAILAGKLLALVCISVVPVLLACGFYFYMALQOFOSRSKLAEESANFACETSIRT 960
Db      901 IITAILAGKLLALVCISVVPVLLACGFYFYMALQOFOSRSKLAEESANFACETSIRT 960
Qy      961 VASLIRERWELIYHAQDLOAGRTSLISYLRSSLVLAASSALVFCVVALGFWGCTLLGH 1020
Db      961 VASLIRERWELIYHAQDLOAGRTSLISYLRSSLVLAASSALVFCVVALGFWGCTLLGH 1020
Qy      1021 HEYDIFRFVFCSEILFGAOSAGTVSFAPDMGKAKNAAAEFRRLDRKPOIDNMSSEGE 1080
Db      1021 HEYDIFRFVFCSEILFGAOSAGTVSFAPDMGKAKNAAAEFRRLDRKPOIDNMSSEGE 1080
Qy      1081 KLETVGEIEFRNVHRPRTPEQPYLRGLDLTVKPGOVYALVGPSCGCKSTTIALLERP 1140
Db      1081 KLETVGEIEFRNVHRPRTPEQPYLRGLDLTVKPGOVYALVGPSCGCKSTTIALLERP 1140
Qy      1141 YDAIAGSILVDGDKISKLINSTRSLVSOEPTLYOGTIKENIILGIVEDVPEBEFLI 1200
Db      1141 YDAIAGSILVDGDKISKLINSTRSLVSOEPTLYOGTIKENIILGIVEDVPEBEFLI 1200

```

```

Qy      1201 KACKDANIYDFJMSLPEGFVTVSGKGMISGOKORVALARALLRDPKILLDEATSAI 1260
Db      1201 KACKDANIYDFJMSLPEGFVTVSGKGMISGOKORVALARALLRDPKILLDEATSAI 1260
Qy      1261 DSESEKVOAALDAARGRTTIAVAHRLSTIQADVIYVFDQGIYESGTHSELVOKGR 1320
Db      1261 DSESEKVOAALDAARGRTTIAVAHRLSTIQADVIYVFDQGIYESGTHSELVOKGR 1320
Qy      1321 YVELVNLQSLGKH 1334
Db      1321 YVELVNLQSLGKH 1334

RESULT 3
US-08-612-734B-2
; Sequence 2, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center, DC1501
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-734B-2

Query Match      80.5%; Score 5407.5; DB 2; Length 1349;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 1069; Conservative 114; Mismatches 141; Indels 27; Gaps 6;

Qy      5 ETNPLSPETMARPAETSTTEEOASPHADEKKILSDLSAPSTWATPPADKEHPRSS 64
Db      5 ETGASSREKSL-EDLYATLETKRSTSSGADNEKPRDHSLSDTINAPD-----GKKRD 59
Qy      65 SNNAVSYNEVDALIAHLPEDEROVLKTOLKEIKVNIISFGIMRIVATKMDILIMVISTICA 124
Db      60 HGRAVDLND-DSLFAHLOEHEKEVILKROLDAPVSVSFFTLRYASRKDILLILVSAICA 118
Qy      125 IAA-----ASTFORIMLYOISYDEFDYELTKNLYEYLLIGEFVYVYVSGEITYVYST 170
Db      119 IAAGAALPFTLILGSLASFQGISGTMPYHEFYHRLTNVLYFYLLGIAEEVTVYVST 178
Qy      171 VGEITYGEHAATOKIREYLESILRONIGYFDKLGAEVTTTITADTNLIQDGISEKVGTL 230

```

```

|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 179 VEFITGEHLTKIRENYEALIRKMMAYFDKIGAGEVTRITADNLQDALSEVGLT 238
OY 231 LVALATFVAFTIAYKYKWKALICSSITVALVLTGSGSOFIIRKSKSIDSYGAGTY 290
Db 239 LFAFATFVAFTIAYKYKWKALICSTIVALVMWGGSRFIVKSKSISEYAGAGTY 298
OY 291 AEEVJSSINATAFGTQDKAKOYEYHDLDAEKWGTKNQIWMGFMIGAMGLMYSNYGLG 350
Db 299 AEEVJSSINATAFGTQDKAKOYEYHDLDAEKWGTKNQIWMGFMIGAMGLMYSNYGLG 358
OY 351 FMMGSRFLVDGAVDGDITVLMALIGSFISGNVSPNAOAFNNAVAAAKIFGTDRS 410
Db 359 FMMGSRFLVDGAVDGDITVLMALIGSFISGNVSPNAOAFNNAVAAAKIFGTDRS 418
OY 411 PLDPYSNEKTLDFEGHIELRNKHIYPSREVTIMEDVSLSPMAGKTALVGPSSGK 470
Db 419 PLDPYSNEKTLDFEGHIELRNKHIYPSREVTIMEDVSLSPMAGKTALVGPSSGK 478
OY 471 STVVGIVEFYPYVKTGVLDDHDKDLRLRQOISVSOBPVLEGTIYKINRHGLI 530
Db 479 STVVGIVEFYPYVKTGVLDDHDKDLRLRQOISVSOBPVLEGTIYKINRHGLI 538
OY 531 GTRKENSEDKYRELJENAKKANADFTALPEGYETVVGORGFLLSGGOKORAIARA 590
Db 539 GTRKENSEDKYRELJENAKKANADFTALPEGYETVVGORGFLLSGGOKORAIARA 598
OY 591 VVSDPKILLDBATSALDTKSEGVVQALERAEGRTIYIAHRTSTIHANIVLVNG 650
Db 599 IYSDPKILLDBATSALDTKSEGVVQALERAEGRTIYIAHRTSTIHANIVLVNG 658
OY 651 KTAEGTDELVDKGAIVKLYEAOIRINOKRDALE--DADAEDLVIADIAKTASS 707
Db 659 KTAEGTDELVDKGAIVKLYEAOIRINOKRDALE--DADAEDLVIADIAKTASS 718
OY 708 ASSDDG---KPTTIDRTGTHKSVSALISKRPPETTPKYSILMTLKRVASPNREIY 763
Db 719 SSNSLIDAVDERAKLEKRRGTOKSVSSAVLSKKNVEQFKISIMLVKIGAFNRRELEY 778
OY 764 MLIGLVSVLGGGQPTQAVLYAKAISTLSLPESOYSKLRHDADEFWLSMFVVGIIQFT 823
Db 779 MLIGLVSVLGGGQPTQAVLYAKAISTLSLPESMHKLKRDHAPFSLMFEVVGIAQFIS 838
OY 824 OSTNGAFAVCSERLIRARSTAFRIILRODIAFDKENSNGALTSFSTETKHLSGVS 883
Db 839 IJSTNGAFAVCSERLIRARSTAFRIILRODIAFDKENSNGALTSFSTETKHLSGVS 898
OY 884 GVTLLGTLTSTTGAALITLALIGMKLALVCISVVPVLLAGGFYFVYMAOFOSRSKLA 943
Db 899 GVTLLGTLTSTTGAALITLALIGMKLALVCISVVPVLLAGGFYFVYMAOFOSRSKLA 958
OY 944 YEGSANFACATSSIRTVASLTREDVWEIYHAQIDAQRTSLISVLRSSLLYASSQALY 1003
Db 959 YEGSANFACATSSIRTVASLTREDVWEIYHAQIDAQRTSLISVLRSSLLYASSQALY 1018
OY 1004 FFCVALLGFYGGTLLGHHYDIFREFVCSSELLFGAQSAGYFAPRPMGKAKNAAEFR 1063
Db 1019 FFCVALLGFYGGTLLGHHYDIFREFVCSSELLFGAQSAGYFAPRPMGKAKNAAEFR 1078
OY 1064 RLFDDRPOIDNNSSEGEKLEYEGELEFRFNHFRKPTRRQCPVLRGLDITVAPGOVALY 1123
Db 1079 RLFDDRPOIDNNSSEGEKLEYEGELEFRFNHFRKPTRRQCPVLRGLDITVAPGOVALY 1138
OY 1124 GBSGGKSTTALLERFYDAIAGSILVDKDISKINSYRSFSLVSQEPVLYOGTKE 1183
Db 1139 GBSGGKSTTALLERFYDAIAGSILVDKDISKINSYRSFSLVSQEPVLYOGTKE 1198
OY 1184 NILLOIVEDDVPEFLIACADANTYDFIMSLPBPNTVYVSGKGLSGGOKORAIARA 1243
Db 1199 NILLOIVEDDVPEFLIACADANTYDFIMSLPBPNTVYVSGKGLSGGOKORAIARA 1258
OY 1244 LLRDKKILLDBATSALDSESEKRVQOALDAARGRTTIAVHRTSTIOKALVITVFPQG 1303

```

```

Db 1259 LLRDKKILLDBATSALDSESEKRVQOALDAARGRTTIAVHRTSTIONADITVFPQG 1318
OY 1304 KIVESGTHSELVOKRGYYELVNLQSLGKH 1334
Db 1319 KIVESGTHSELVOKRGYYELVNLQSLGKH 1349

RESULT 4
US-08-395-246C-2
; Sequence 2, Application US/08395246C
; Patent No. 5773214
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
; TITLE OF INVENTION: ASPERGILLUS FLAVUS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,246C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35784
; REFERENCE/DOCKET NUMBER: x9683
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; TELEFAX: 317-277-1917
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1307 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-395-246C-2

Query Match 59.4%; Score 3992; DB 1; Length 1307;
Best Local Similarity 60.2%; Pred. No. 0;
Matches 790; Conservative 211; Mismatches 244; Indels 68; Gaps 11.

OY 60 PRSSNNNAVSNEVDALIAHPEDEROYLKTQLEIKVNFSGFGLMGRATKMDILIMYI 119
Db 13 PRSGPSTTT-----GHSVSHAEEVLDRLQHLTPVSOISFQIYRATRMWDAILLG 64
OY 120 STICAIA-----ASTFORIMLYOISYDEYDELTKNVLYFYVYLGIEFVT 165
Db 65 SALAIAAGALPLFTVLEGRLTSTFODIATHRIYDHHHELTNNVYFYLGAAEVYA 124
OY 166 VVVSIVGPIYGEHAOTKIREYLSILRONIGYFDKLAGAGVTRITADNLIDGISE 225
Db 125 IYLAIVGFIYTDHVOQIRVEYFOAILLQNIATFDTLGAEGITRTITADTNLIIDGISE 184
OY 226 KYGLTLTAFVTAFTIAYVYWKALICSTIYVALVLTGSGSOFIIRKSKSIDSYG 285
Db 185 KYGLTLTAFVTAFTIAYVYWKALICSTIYVALVLTGSGSOFIIRKSKSIDSYG 244
OY 286 AGGTAAEEYSSIRNATAFGTQDKAKOYEYHDLDAEKWGTKNQIWMGFMIGAMGLMYS 345
Db 245 RGSMAEDILDSIRTVAFNMAOETIARKYESHLKDAEGPKASKYFIYFIMGALCITYL 304
OY 346 NYGLGFMGSRFLVDGA--VDVGDIITVLMALIGSFISGNVSPNAOAFNNAVAAAKIF 403

```

```

DB 305 NGLGFMWGSRLVEGISNFKADVLTIMAAIILGSSYNLGNVAPNQAALSDAFAAASKLY 364
QY 404 GTIDROSPLDPVSNKGLTDHFEHIELRNVKHIYPSRPEVTWEDVYLSMPGKTTALV 463
DB 365 GTIDROSPLDASDQKTYLEFVNGNLYLVQIRVYPSRPEVYAHNDISCTYIPGKTTAFV 424
QY 464 GPGSGSKTVGLVEFVFPVGRVTLLDGHIDKDLRWLRQOISLVQSPVFGTTIK 523
DB 425 GPGSGSKSTIILERFYPVAVATIMLDGIDITLNRWRQOMSLVQSPRIEFTATIAE 484
QY 524 NTRHGLIKTYENESDQVRELIENAKANAHDFITALEPEYETVNGRGFTLSGQKO 583
DB 485 NTRGIGISRFKESTYEIRKREAAARAHADFTALPNGYDTN--ESFISGQOKO 542
QY 584 RIAIAVAVSDPKIILDEATSLDTRKSEGVQVAALERAEGRTTYIARHLSITKTAH 643
DB 543 RIAIAARAILDKPILLDDEATSLDTRKSEGVQVAALERAEGRTTYIARHLSITKTAH 602
QY 644 IYLVNGKIAEQTHDELVDYRGAYRKLVEAQRINEQKEDALEDAEDLTN----- 696
DB 603 IYLVNGKIAEQTHDELVDYRGAYRKLVEAQRINEQKEDALEDAEDLTN----- 660
QY 697 -----ADIATKTASSASDDGKPTTIDRTGTH-----KVSASALSK-R 736
DB 661 MPEFPKDYPGDESDIYSI--LSDNASDI-----GLHTGEKORVPSRMSLSHM 708
QY 737 PREPTKYSIWLTLKFPASNPRIPIYMLGLVPSVLAGGPOPAVLYKAIISTLSPE 796
DB 709 QPVEKENASVWTLKFKPLASNPRIPIYMLGLVPSVLAGGPOPAVLYKAIISTLSPE 768
QY 797 SOYSKLRHADFWLMEFVVGIIQITOSTNGAFAVCSRLIRARSTAFRTILRODIA 856
DB 769 LEYKLRHADFWLMEFVVGIIQITOSTNGAFAVCSRLIRARSTAFRTILRODIA 828
QY 857 FPDKENSTGALTSTSTETKHLGVSQVTLGTLMTSTLGAIIITAIAGWKALVCI 916
DB 829 FPDQOEWTLGALTSTSTETKHLGVSQVTLGTLMTSTLGAIIITAIAGWKALVCI 888
QY 917 SVYPRVLLAGCFVRYMAOQOSRSKLAEGSANPACGATSIPIRVYALTERDQWELVYH 976
DB 889 SVYPRVLLAGCFVRYMAOQOSRSKLAEGSANPACGATSIPIRVYALTERDQWELVYH 948
QY 977 QLDAGQSTLSIYLRSSILYASSOALVFCVVALGFVGTLLGHHEYDIRFEVCFSEIL 1036
DB 949 QLDAGQSTLSIYLRSSILYASSOALVFCVVALGFVGTLLGHHEYDIRFEVCFSEIL 1008
QY 1037 FGQOSAGTVSFAPDMGKAKANAARERLDRKQIDNMSBEQKLETVEGEIEFRNVHF 1096
DB 1009 FGQOAGTVSFAPDMGKAKANAARERLDRKQIDNMSBEQKLETVEGEIEFRNVHF 1067
QY 1097 RYPTPRPOVPLRGDLTVKRGQVAVYVAGPSCGKSTTITALLERDYDAIAGSLVDYDIS 1156
DB 1068 RYPTPRPOVPLRGDLTVKRGQVAVYVAGPSCGKSTTITALLERDYDAIAGSLVDYDIS 1127
QY 1157 KLINISYRSLVSOEPRTLYOGTIKENILGLIYEDVPEEFILKACKDANIYDFIMSLP 1216
DB 1128 KLINISYRSLVSOEPRTLYOGTIKENILGLIYEDVPEEFILKACKDANIYDFIMSLP 1187
QY 1217 EGRTVTVGSGKMLSGQOKORVATARALLDPKILLDDEATSLDSESEKVVQALDAAA 1276
DB 1188 EGRTVTVGSGKMLSGQOKORVATARALLDPKILLDDEATSLDSESEKVVQALDAAA 1247
QY 1277 RGRTTIVARHLSITOKADVIYVFDQKIYESGHSSELVYOKKGRYELVJNLS 1329
DB 1248 RGRTTIVARHLSITOKADVIYVFDQKIYESGHSSELVYOKKGRYELVJNLS 1300

```

RESULT 5
 US-08-612-521-2
 ; Sequence 2, Application US/08612521
 ; Patent No. 5786463
 ; GENERAL INFORMATION:

```

? APPLICANT: Peery, Robert B
? APPLICANT: Skatrud, Paul L
? APPLICANT: Thornevell, Susan J
? TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
? TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
? NUMBER OF SEQUENCES: 3
? CORRESPONDENCE ADDRESS:
? ADDRESS: Patent Division/AEH
? STREET: Lilly Corporate Center
? City: Indianapolis
? STATE: Indiana
? COUNTRY: USA
? ZIP: 46285
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/612,521
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Hamilton, Amy E
? REGISTRATION NUMBER: 33,894
? REFERENCE/DOCKET NUMBER: X-9693
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 317-276-3169
? TELEFAX: 317-276-1294
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1408 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-612-521-2

```

```

Query Match 39.9%; Score 2682.5; DB 1; Length 1408;
Best Local Similarity 43.9%; Pred. No. 2.6e-234;
Matches 597; Conservative 233; Mismatches 449; Indels 81; Gaps 16;

QY 33 AADKKILSLDLSAPSSGTTA--TPADKEHRPKSS--SSNNVAVSNEVDALLIAHLPEDEROVL 89
DB 69 AAEKIRSSSIASPSDLRMSPLEK---PISNAFSKSHPYKSKFDFLKSRRKKEEBER 125
QY 90 KTOLE---IKVNSFEGIMRYATKMDILMIVISTGAIASA-----T 130
DB 126 NKEKEKASVLPVPSFALFEFAPLETIAVNLGLVLAAGSCQPLMTLITFGRLTTSFT 185
QY 131 FORIMLYOISY-----DEFYDELTKNLYEYVLGIGEFVTVYVSTVGFIY 175
DB 186 NYAVIANQISGGGLTPETSAAALQAADKLTQSGHNALYIMAAIGMELAWLWLMVIMNV 245
QY 176 TGEHATOKIRIYTESILRQNTGIFDKLIGAEVTRITADTNLQDGISEVGLITLALA 235
DB 246 TGEHNSKRIRERLYAAALROEIAFDDLAGAEVARIOTDCHLVQEGSEKVALVFOYAG 305
QY 236 TEVAFATIAVYKWKLLICSTIVAVLVYMGGGSQFITYKSKSLDSYGAGTVAEEVI 295
DB 306 TVVCGFVIAFVRSRLGALVSIIPVIMLCGIMMTAMAKGTIALDHIAAGSLAEEVI 365
QY 296 SSINATAFTQDRLANQOYEVLHDEAEKMGTKNOIYMGFMIGAMFGMTSYNYGLFWMS 355
DB 366 GSITVQAFGEKELIGKFAHIDQSKLVGRKSGIFBGFGLSIMEFVYIAVYALAFYGG 425
QY 356 RFLVDGAVDGDILITLMAIILGFSGLGNSVPAQAFTNNAVAAAKFGITIDROSPLDPY 415
DB 426 ILVNSGQADSGIYINVMKSLIGFSFSAKMLAPBELAAVYKAGAAKLFATIDRPALDSA 485
QY 416 SNEKTLIDHFEHIELRNVKHIYPSRPEVTWEDVYLSMPGKTTALVAGPSGSKSTVYG 475
DB 486 SEBGFKPDGLNGELISFENVKFHYPSRPSIPILKGTPTTFEAGKTFALVAGSGSKSTVYS 545

```

[illegible]

```

Db 904 FRTVSLTOROKFEHMYAOSLQVYRNLSLRKAHIFGITFSFTQAMMFYSGCFRGAYL 963
QY 1018 LGHHEVDIFFEFCSEFIILFGAQSAGTVFSFADMGKAKNAAEFRRLFPKPIDWMS 1077
Db 964 VAHKLMSFEDVLVFSANVFGAMAVGVSSFPADYAKAKISAHIIMIIEKTPIDISYST 1023
QY 1078 EGEKLEVEGEIEFRNHFYPTRPBPQVLRGLDLYKPGQYVALVPGSCGKSTYALL 1137
Db 1024 EGMPTLBEQNTVFGEVVFYPTRPDIPLVQGLSEVKKQOTLALVSSCGKSTYVOLL 1083
QY 1138 ERFYDAAGSILVDGDISKLINSTRSELSVLSQEPFLYQGTIKENILLGIYEDVPEE 1197
Db 1084 ERFYDPLAGVLLDGEIKRLNQLFAHLGIYQSEPIELDCSIAENIAYGDNSSRVSOE 1143
QY 1198 FLKACKDAITYEIMSLPEGEFTYVSGSGGKMSGGOKORAVATARALLRDPKILLDEAT 1257
Db 1144 EIVRAKKAENIHAIFIESLPKYSTKYGDKCTQJSGGOKRIARALRVLPHILLDEAT 1203
QY 1258 SALDSESEKVVQAALDAARGRTTIAVAHRLSTLOKADVIYVEPOCKIVESGTHSELVOK 1317
Db 1204 SALDTESEKVVQAEADARGRTCIYAHRLSTIQNADLLIVFPONGVKEHGTHQQLAQ 1263
QY 1318 KGRYELVNIQS 1329
Db 1264 KGIYFSMVSVOA 1275

RESULT 7
US-08-583-276-19
Sequence 19, Application US/08583276
Patent No. 5837536
GENERAL INFORMATION:
APPLICANT: McDonagh, Kevin T.
APPLICANT: Nienhuis, Arthur
TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Giffillan,
ADDRESSEE: Cecchi & Stewart
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: DM4.V2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,276
FILING DATE: 05-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/332,444
FILING DATE: 31-OCT-1994
APPLICATION NUMBER: 07/887,712
FILING DATE: 22-MAY-1992
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1280 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
US-08-583-276-19
Query Match 36.9%; Score 2477; DB 2; Length 1280;

```

```

Best Local Similarity 41.9%; Pred. No. 1,le-215;
Matches 533; Conservative 230; Mismatches 449; Indels 60; Gaps 9;
QY 94 EIVVNISFGLMRYAKRIMOLLIVISTICIAAASIFQRIMLY----- 137
Db 28 KKKPFTVSVSMFYSMWMLKLYMVGSTLAIIGAGLPLMLLFGENTDIFANAGNLED 87
QY 138 -----QISYEYF----DELTKNLYFVYIGIEFVYVSVTGFTYGGHATQKI 184
Db 88 LMSNITRSDINDGFMNLEEDMTTRAYYSIGAGVLAAYQVSWCLAAARQIHKI 147
QY 185 REYVLESILRONIGYFDKLGAGEVYTRITADTNLIQDISERKVLPTLALAPVTAFLIA 244
Db 148 RKQFHAIMRQEIQMFVHDVGEINTRLTDVSKINEVICDKIMFQSMATFFGTGIVG 207
QY 245 YVTKWKLALICSSITVALYLMGGSGQFIYKSKSDSYAGGTVAAEEVSSIRNMTAF 304
Db 208 FTRQWKLTLYLALSPYLIGLSAAVWAKILSFTDKELIAYAKAAVAEEVLAALRYIAF 267
QY 305 GTQDKLAKQYEVHLDEAEKMGTKNQIYWGEMIGAMFGIMYSNYGLGEMSGRPLVDGAVD 364
Db 268 GGQKKLELRKYNKLEAKRIGIKKAITANISIGAAFLIYASALAFYGTTLVSGEYS 327
QY 365 VGDILYVLMALIGSFSLGNVSPNAQAFITNAVAAAKIEGTIDROSLDPYNSGKTLDH 424
Db 328 IGVLTVEFFSVLIGAFSGQASPSIEAFANARCAAYEIFKIIDNKPSIDSYSGHKPDN 387
QY 425 FECHIELRNKHYTPSREPVYVMDVSLSMRAGKTALVYSGSGSKSVYGLVERFYVY 484
Db 388 IKGNIEFRNHFVSYSRREVRILKGLNLYKQSGQTVLVAGNSGGKSTVQIMORLXDPT 447
QY 485 RGTVLDOHDKDINLRLKROQISLVSOEPLFGTTYIKNRHGLIGTKYENESDKYRE 544
Db 448 EGMYSVGDQDITRNRLREIIGVSOEPLFATTIENIRYG-----RQNTYDE---- 499
QY 545 LIENAKMANAHDFTALPEGYETNVGGRFLSGGOKORIAARAVYSDPKILLDEAT 604
Db 500 -IEKAVKANAYDTIMKPHFEDLVGERGALSGGOKORIAARALVRNPKILLDEAT 558
QY 605 SALDTSRGVQAALERAEGRTIYVAHRLSTIKTHNIYVNVNGKIAEGTHDEVD 664
Db 559 SALDTESEAVVQALDKARKRTTIVAHRLSTVRNADVIAGFDQVVEGNDHELMKE 618
QY 665 GCAVRKIVEAQRINDQKADALEADLNADIAKIKTASSASDDIGKPTTIDRTGT 724
Db 619 KGIYFKLVTQTAGNEVE---LENA-----ADESKEDIDALEKSSDSSSLIRKST 668
QY 725 HKVSSAALSKRP-----PETPKYSLWTLKRFVASFNPPEIPYMLIGLVSVYLAGG 777
Db 669 RRSYRGSAQDRKLTSTKEALDESIPVPSFWIRIMK---LNTLEMPYFVYGVFCALINGL 724
QY 778 OPTQAVLYAKAISTLSPEOSYKLRHDQFMSLMFFVGIIOFTOSTNGAFAVCSER 837
Db 725 QPAPALISKLIIGVTRIDDEPK-RQNSNLSFLFALGISFTFLQGFTEGKAGEI 783
QY 838 LIRRASTAFPTIIRODIAFPDKEENSGALTSFTFKHLSGVSYTLITLMTSTYL 897
Db 784 LTKRLRYVFRSMLRQDYSWDDPKNTGALTTRILANDAAQYKAGISRLAVITQNIANTL 843
QY 898 GAAIITIALIGWKIALVCSIVPVYLLACGYRFYMLAOFQSRSKLAYBGSANFACEATSS 957
Db 844 GTGIISFIYGMQTLILLALIVPIIAGVEMKLSGQALKDKKELGAGKAIATEIEN 903
QY 958 IRTVASLTRDQWMEIYNAQDQAGRTSLIVRSLLIYASSQALYFCVALGFMYGITL 1017
Db 904 FRTVSLTOROKFEHMYAOSLQVYRNLSLRKAHIFGITFSFTQAMMFYSGCFRGAYL 963
QY 1018 LGHHEVDIFFEFCSEFIILFGAQSAGTVFSFADMGKAKNAAEFRRLFPKPIDWMS 1077
Db 964 VAHKLMSFEDVLVFSANVFGAMAVGVSSFPADYAKAKISAHIIMIIEKTPIDISYST 1023
QY 1078 EGEKLEVEGEIEFRNHFYPTRPBPQVLRGLDLYKPGQYVALVPGSCGKSTYALL 1137

```


Db 1024 EGIAMPYTLGNTFGEVENVPTPRDIPVLOGLSLEVKKGQTLALVSSGCGKSTVVOQL 1083
 Qy 1138 ERFYDAIAGSLVDGDISKLININSYRSLSVSOEPTLYOGTIKENILLGIVEDVPEE 1197
 Db 1084 ERFYDPLAKVLLDGKEIRLNVQMLRAHLGIYSQEPILFDCSIAMENIAYGDSRVVSOE 1143
 Qy 1198 FLIKACKDANIYDFIMSLPEGEFTVVSCKGMSGQOKORVAIARALLDPKILLDEAT 1257
 Db 1144 EIVRAKENIHAFIESLNPSTKYDGKGTQLSGQOKRIARALVROPHILLDEAT 1203
 Qy 1258 SALDSESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVDOCKIYESGTHSELVOK 1317
 Db 1204 SALDTESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVDOCKIYESGTHSELVOK 1263
 Qy 1318 KGRYYELVNLQS 1329
 Db 1264 KGIYFSMVSVOA 1275

RESULT 8

US-08-784-649A-2
 ; Sequence 2, Application US/08784649A
 ; Patent No. 5830697

GENERAL INFORMATION:

; APPLICANT: SIKIC, Branimir I
 ; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
 ; TITLE OF INVENTION: CYCLOSPORIN MODULATION
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,649A
 ; FILING DATE:

CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sherwood, Pamela J
 ; REGISTRATION NUMBER: Reg. No. 5830697 36,677
 ; REFERENCE/DOCKET NUMBER: 06037/007001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-322-5070
 ; TELEFAX: 415-854-0875
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1279 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-784-649A-2

Query Match 36.7%; Score 2469.5; DB 2; Length 1279;
 Best Local Similarity 42.0%; Pred. No. 5.3e-215;
 Matches 534; Conservative 229; Mismatches 448; Indels 61; Gaps 10;

Qy 94 EEIKVNIISPEGLMRYATKMDIIMVSTICATAAATFORIMLY----- 137
 Db 28 KKKKPVSVFSSMRYSNMIDKLIMVGTAAIIGHAGLPLMLVGEEMTDIRANAGNLED 87
 Qy 138 -----QISYDEY-----DELTKNVLVYFYLLGIGEEVTVVSTVGFITYGEHATOKI 184
 Db 88 LMSNTINRSDINDTGFPMLEEDMTIRAYAYYSIGIGAGVAVAYIYQVFCIAGAAGROIHKI 147

Qy 185 REYELSLIKONIGYEDKLGAGEVTTTRITADNLIODGISSEKVGTLTALAEFTAFITIA 244
 Db 148 RKOFPAHNMQOEIGWPDVHDVGEELNTRLDVDSKINEVEDIKGMFGQSAFFTEGFIY 207
 Qy 245 YVYKRWALICSSITVALVITMGSSQFIITKYSKSLDVGAGGYVAEEYISSIRNATF 304
 Db 208 FTRGMKTLTYVITLAIISPVLGISAAMVMAKILSSFFDKELLVAKGAAVEELAIRVIAF 267
 Qy 305 GTQDLAKOYEVALLDEAEKMGTKNQIYMGFMIGAMGLMYSNGLFMGMSRFLYGAAD 364
 Db 268 GGGKTELEENKMLEAKRIGIKKAITANISIGAAFLITIASALAFVGTITVLSGEYS 327
 Qy 365 VEDILFVLAAILIGSFLGNSVSPNAOFTYNAVAAAKIFGTIDRQSPLDIPYSNEGTLIH 424
 Db 328 IGGVLTIV-FSVLLIGAFSVGASPSIEAFANRGAAYEIFFIIDKNPSIDYSKSGHKPDN 386
 Qy 425 FEGHIELRWKHIYPSRPETVMEVDSLSMPAKTALVSGSSGKSTVVGLEYERYMYV 484
 Db 387 IKGNLEFRNVHVSYPKRKEVILKGLNLVQSGQVAVLGVNSGCGKSTVVOQLRLYDPT 446
 Qy 485 RGTVLLDGHDIKDNLNRMRLROQISLVSOEPLVLEGTIYKIRHGLTGYENESDEKVE 544
 Db 447 EGMVSVDOGDIRTINRFLREIIGVVSQEPVLEFATITAIENIRG-----RENTYME---- 498
 Qy 545 LIEAANKANAMADFTALPEGEFTVNVGQCEFLISGQOKORIAIARAVSDPKILLDEAT 604
 Db 499 -IEKAVKENAVDYFIKLPKFPDLVGERGAQLSGQOKORIAIARALVNPDKILLDEAT 557
 Qy 605 SALDTESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVDOCKIYESGTHSELVOK 664
 Db 558 SALDTESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVDOCKIYESGTHSELVOK 617
 Qy 665 GGAYRKLVEAORINQEKADDAEDLTNADIKITASSASSDLDGKPTTIDRTGT 724
 Db 618 KGIYFLVYTMQTAGNEVE--LENA-----ADESKSEIDALEMSNDSRSSLLIKRST 667
 Qy 725 HKVSSAIIISKRP-----PETTPKYSLMTLLKFAVASTNRPEIPMLIGLVSVIAGCG 777
 Db 668 RRSVRGSOAQODRLSTYKALDESIPVPSFWRIK---INLEMPYFVGVFCAIINGSL 723
 Qy 778 OPTQAVLYKKAISTLSLPSQVSKLRHDDFMSLFEVVGIIOTIQSTNGAFAVNCSSR 837
 Db 724 QPFAIIFRSKIIGVFTRIDDPETK-RQNSNLSLELALGIISFTITPFGFTFGAGEL 782
 Qy 838 LIRARSTAFRILRODIAFFDEKENSATGALSTFETKHLSGVSGVLTGTLMTSTTL 897
 Db 783 LTRKRLYVFRSMLRQDVSFDPKNTGALTRLRNLDAAYKAGALGSLAVITQIANL 842
 Qy 898 GAATITLALIGKMLAVCSIVPVVLLACGFYRPMYLAOFOSRSKLAYESANFACATSS 957
 Db 843 GTGIIISFTYGMQTLTLLLAIVPIIAGVEMKMSGQALDKKELBAGKATATEAIE 902
 Qy 958 IRTVASLTEREDVWEIYHQAOLDAQGRSTLSIVRSLSLVAASSOALVFCVAGFWGTL 1017
 Db 903 FRTVVSLOTQKFEHMYTASLOVPYRNSLRKAHIFGITTSFTQAMMYFSYACGFRGAVL 962
 Qy 1018 LGHHEYDIFRFYCFSEILFGAOSAGTIVSFAPDMKANNAAEFRLDRKPOLIDMSE 1077
 Db 963 VAHKLMSEFDVLLVFSAYVAGAAVGVSSFAVDYAKAKISAHIMIIETKPLIDSYST 1022
 Qy 1078 EGKELTVEGEIEFRVNRVRYPRPEQPVLRGIDLTVKQGVYALVPSGCGKSTIALL 1137
 Db 1023 EGIAMPYTLGNTFGEVENVPTPRDIPVLOGLSLEVKKGQTLALVSSGCGKSTVVOQL 1082
 Qy 1138 ERFYDAIAGSLVDGDISKLININSYRSLSVSOEPTLYOGTIKENILLGIVEDVPEE 1197
 Db 1083 ERFYDPLAKVLLDGKEIRLNVQMLRAHLGIYSQEPILFDCSIAMENIAYGDSRVVSOE 1142
 Qy 1198 FLIKACKDANIYDFIMSLPEGEFTVVSCKGMSGQOKORVAIARALLDPKILLDEAT 1257
 Db 1143 EIVRAKENIHAFIESLNPSTKYDGKGTQLSGQOKRIARALVROPHILLDEAT 1202
 Qy 1258 SALDSESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVDOCKIYESGTHSELVOK 1317


```

: SEQ ID NO 2
: LENGTH: 1275
: TYPE: PRF
: ORGANISM: HOMO SAPIENS
US-09-450-105-2

Query Match      35.8%; Score 2408.5; DB 4; Length 1275;
Best Local Similarity 40.7%; Pred. No. 1.9e-209;
Matches 515; Conservative 231; Mismatches 463; Indels 57; Gaps 8

QY      94 EEIKVNISFEGIMRYATKMDILINVISITCAIAASTFORIM----- 136
Db      26 KEKKPAVGIQGFMRADWIDKLCMALGTIAIHTGLPLMLVGYMTDSFTQAEIRLL 85
QY      137 -----YQSYDEYFDELTKNLYEYVYLGEEVYVYVSVGVYIYGEHATQKI 184
Db      86 PSYTNQSEINSTQYVSDSSLEEDMAMAYAYYYGIGAGVIAIVAYIOVSLMCLAAGRQIKHI 145
QY      185 REYLESILRONIGYGFDRKIGEVTRTIRADPNLIDQDISEYVGLTFLAFATFVPIIA 244
Db      146 RQKFPHALINQDEIGTFVDVNDAGELWTRLLRDLVDYSKINDSGDLKGMFPQSITTFSSAGFIIG 205
QY      245 YKKYKLLALICSSITVALVLTIMGSSQFLIKYSKSLDSYGAGGVIAEEVVISIRNATAF 304
Db      206 FLSGKRLTFLIVLASPLGLSLSAMMAKVLTSPTNKELQYAKAGAAVEHVALIRPIVIAF 265
QY      305 GTQDKLAQYEVHLDDEAEKMGTKNOIVMGFMILGAFMGFLMYSIGLGFMMGSRPLVDGAND 364
Db      266 GQKKLELERYNNKMLEAKRVGIKKATITANISIGIYLLVYASYALAFWYGTSLVLSNEXS 325
QY      365 VGDILTFLMALIGSEFSICGNVSPNAQAFNNAAAKIGTIDROSLDPLDPSNEGKTLDH 424
Db      326 IGVQVTLFFSILGTFSGHGLAPNIEAFRNANGAAVEIKITIDNESIDISFTKGKPKDS 385
QY      425 FEGHLELRNVKHIYPSREPVYMEVDSISMPAGKTTALVSPSGSGSKTVVGLVEREYMEV 484
Db      386 IMGNEFERNVYPNYPSRSBVKILKGLINLKVKSGQYALVGNSGCGKSTVYQLQRLRYDPI 445
QY      485 RGTVLIDGHDIKDLWRMLRQOISLVSQEPVLFGTITYNIRHGLIGTYENESDEKVAE 544
Db      446 EGEVSIIDQDQDITINVRYLRELIYGVSGQEPVLFATTAENIRYQ-----RENTYMD----- 497
QY      545 LIENAKKANANDFTTALPEGEYETNVGSGEFLSSGQXORIMARAVASDPRKLLIDEAT 604
Db      498 -IEKAVYKEANADFTLMPKLPKFNNTLVGEAGDLSGGQKORIMARLAVLNPKLLIDEAT 556
QY      605 SALDTRKSEGVQALATERAEGRTTIVIAHRTSTIKTAHNIIVLVNKGIAEOGTHDELVLR 664
Db      557 SALDTESEVAVQALDKAREGRTTIVIAHRLSTVRNADVIAGPDGVIVYEQGNHELMKE 616
QY      665 GGAYKKLYEAO-RINE---QKDALEDADAEDLTNADIAKIKTASASSSDLGKPTTI 719
Db      617 KGIYKRLVMTQRGNEIEEGNNAYESOSPSTGASELTSEKSKSPLIRRSIRRSIHRQDPE 676
QY      720 DRTGHRKSGSSALILSKRPETTPKYSIMLTKLFVASFNPDELPYMLIGIVFSVLACGGOP 779
Db      677 RLKSKSEVD-----EDVPKVSFWOLK-----LNTSEMPYLVGVLCATINCIP 723
QY      780 TQAVLYAKAISTLSPESQSKLRHADAFWMSLMEFVAGIIOETIYSTONAAEAVCSERLI 839
Db      724 VPAIVFSKIVGVFSDDDHETKQR-NCNLFSLFLVMGMISVTTYPFQGFTRKAKIELTL 782
QY      840 RRAARSTAEITILRODIAFFDKENSTGALTSLSTETKHLSGVSGVLTGTLTMTSTTLGA 899
Db      783 KRLRYMFEKSMRLRODISWPDHDKNTTGLSTRILASDASNVKMGMSRLAVVQONVANLT 842
QY      900 AIIYALAGMKALYCISVVPVYLLACGFRFTMLAOFQSRSKYALGCSANFACEATSIIR 959
Db      843 GIILSLVYGWQJTLTLLVYIPLIVIGIITEMILISQALKDKRELEISGRIATEALENR 902
QY      960 TVASLTREDDWEIYHAQDDAGRTSLISLVASSLSQALVPCVALLGFYGGTGLG 1019
Db      903 TVVSLTRDRQKFEETMAQSQIDPIYRNALKKAHVGLTFAFQYMAIFSYACPFEGVAYVA 962

```

```

QY      152 VVFFVYLGIGEVVYVYVGVGFTYIGENATQKIREYVLESILRNQINGVFDKIGAGEVYTR 211
      :|||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
      :|||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Db      963 RELMTFENVMLVFSANVVGMAAGMTSSFPADYAAKAYSASHIITIEIKIPIDBSYTBG 1022
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      1080 EKLEVEGEIEFRNVHFRYPTRPEQPVLRGLDLYVKKPGQYVALVGPSCGKSTTALLER 1139
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1023 LKPNMLFENSVKFNQWFMVNPTRPRNPVPIVGLTSFEWKKCOTLRVSGSSCGKSTVQLLER 1082
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      1140 FYDATAGSLVNDKRIKSLTNINSYSLFSLVSOEPTLVGSTIKENIILGIYDVDPPEPTL 1199
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1083 FYNPAGVIFLDGKEIKIKOLNVOMLAAHGLVISOEPLTFDCSTITENIAYGDSNRVVSHEI 1142
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      1200 IKACDANIYDFIMEILPEGFNTVVGSKGMLSGOKORAYARALLRPKILLDEATSA 1259
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1143 VYAAEENINHGIFDLSPEKYNTRVGDGKQVLGSGGKORAIARALVRPHILLDEATSA 1202
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      1260 LDSESEKVVQALDAANGRTTIAVAHRLSTIQKADVIVFDGKIVBSGTHSELVQKGG 1319
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1203 LDTSESEKVVQAEALDKARREGRCIVIAHRLSTIQKADLVIVIONGVKKEGTHQQLAKGG 1262
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      1320 RYTEL 1325
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1263 IYFSKV 1268
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 13
US-08-996-644-2
Sequence 2, Application US/08996644
Patent No. 5945324
GENERAL INFORMATION:
APPLICANT: Skatrud, Paul L.
APPLICANT: de Waard, Maarten
TITLE OF INVENTION: Multiple Drug Resistance Gene atcc of
TITLE OF INVENTION: Aspergillus nidulans
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,644
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11765
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-644-2

Query Match 32.9%, Score 2208.5, DB 2, Length 1308;
Best Local Similarity 40.3%, Pred. No. 3,le-191;
Matches 479; Conservative 223; Mismatches 456; Indels 31; Gaps 10;

```

Db 140 IIFYVYLGIAIRLVSTYTYTLLTYAAYRIYVRIRHAYLKAALSOEVAAYDFGSGGSIAMQ 199
 QY 212 ITADNLIIDGISEKVGTLTALATFATFATIAVYVYKWLALICSSSTIVAIYVTMGGSQ 271
 Db 200 ATSNKLLIOAGASDKIGLIFOGIAAFVTLISLCKMCKWTLTLCICIPATICTTGVAA 259
 QY 272 FIIRYKSKSLDSYGAGGVAAVEVYSSIRNATAFGTODKLAKOYEVHLEDEAKMGTNNQIV 331
 Db 260 VEAGHETRIQIHAQANSFABEGILACVKAHAFGRDLSVRRKDEYLVFAHKKKISPL 319
 QY 332 MGFPMIGAMGLVSNYGLGFWMGSRFLVDGAV-DVGDILTVMAIILGFSFSLGNVSNQAQ 390
 Db 320 LELTSAEYTYIYLGIAFWOGIHMEGRGEIGTAGDITFVTLISVYASINLTLLAPYSI 379
 QY 391 APTNAVAATAKIGTIDROSPLDPSNEGKTLDFEGHIELRNKVIYPSRPEYTVMEDV 450
 Db 380 EFSRAASAAQFLRILDRSEINPYKEGLEPERVLGDVELEVNTFSYPTRRGITVDNF 439
 QY 451 SLSPMAGKTALVPGSGSKSTVVGILVERPYMFGVGTLLDGHDKDLNRLMQOISLV 510
 Db 440 SLKVPAGKYALVGGSGSKSTVIGLIERWYNTSGAIRDGNLISELVNGLRRNRLV 499
 QY 511 SOEPVLFGTIKNTRHGLIGTKYENESDKVLELLENAMKANADFTALDEGETNV 570
 Db 500 QDEPVLEQSVFDNIRYGLGTPMENASREDEQEROEAKLAHYAHEFISELDGDTLL 559
 QY 571 GORGFLISGQKORIAIAVAVSDPKILLDEATSAIDPFSQGVQOALERAEGRTTY 630
 Db 560 GERGGLISGQKORVIAIASVVSQPKVLLDEATSAIDPFAETIYKALDKAEGRTTY 619
 QY 631 IAHRLSTITAHNIVLVNGKIAEGTDELVDGARYKLEAQR--INEOREADLED 668
 Db 620 IAHKLTATIRKADNIYVMSGHIYEOGTHSLAKDGVYAGIVKIQMLAVASAHADVNEE 679
 QY 689 ADAEDLTNMDIKIKKASSASSDLDGKPTTIDRTGHSYSSAILSKRPETPKXSLWT 748
 Db 680 GBEDEDALLEVE-----PAVTPPT--SIRGRMNSIKDRDYENKHM 722
 QY 749 LKLFVASPNR--PEIPYMLIGLIVSVLAGGQ-PTQAVLYAKAISTSLP-ESQYSKLRI 804
 Db 723 MLAAALAIYRBECELKAWILVVLGGLGCGAMYPQAILMSRYVEVFTLSGDMADK--- 779
 QY 805 DADFWSLMFEVVGITQITOSTNGAFAVCSERLIRANSTAFRTLLRODIAFDKEENS 864
 Db 780 -GDYFASMLIVLAACGLICYLAVGATNTIAQHLSHMFRLLIHDMLROIQFDEBENT 838
 QY 865 TGAULTSFLSTETKHLISGVGVTLLTLMSTLGAIIITAIAGKIALYCI-SVAVLL 923
 Db 839 TGAUVSRIDSYPHAILLEMGYNIAVLAVLQVVTGIIAIAFSMKGLVNVFEGGIPPLV 988
 QY 924 ACIFYFYMLAOFOSRSLKIEGASANFACATSSIRTVASLTREPRVWEIYHAQOLAQR 983
 Db 899 GAGMVAIRDSRLDRQTSKRYGTSISSIASRAVAIAITVSSLAEEYVLRRTYELDHAYS 958
 QY 984 TSLISVLRSSLIYASSQALVFECVALGFWMYGTLLGHHEYDIFRFVCSFSEILFGASAG 1043
 Db 959 SSYKPAAMATMIGLTCQIEYMFQALGFWMYGCRLVSGTSMYSFVAFLSVFAGQASA 1018
 QY 1044 TVYSEFADMGKAKANAARERLFDKRPQIDNMBEGBEKELETYGELEFRVHRYPIRPE 1103
 Db 1019 QLFQWSTSTIKGINATNYIAMLHQLOPTYAREPENHDKGGSGAPIAMDVRESYPLRPD 1078
 QY 1104 QPVLRIIDLTVKRGYVAVVAVPGSGCKSTTIALIERFYDIAAGSILVADGDIKLNINSY 1163
 Db 1079 APLIKVNLKINKGPIIAYVGGSGCKSTIMAILERPYDPTGTSITIDATLIDINISY 1138
 QY 1164 RSFLSLVSOEPTLYGCTIKENILGLIVEDVPEEFILKACKDANIYDFINLSDEGEFTV 1223
 Db 1139 RNIVALVQOEPTLFGCTIRINISIGDAVKSVDQIESALRANAMDFVSLPQGITPA 1198
 QY 1224 GSKGMLSGGOKORVIAIARLLRDPKILLDEATSAIDSESEVYQALDAAR--GRTT 1281
 Db 1199 GSGGSLSGGOKORIAIARLLRDPKILLDEATSAIDSESEKIVOKALEGAARDGDRLT 1258

QY 1282 IAVAHRLSTOKADVIYVPEQKIVESGTHSELVOKRGREYELVNLQSL 1330
 Db 1259 VAVAHRLSTIKDANVLCYFEGKIAEMGTHQELIVGGLVRMCEAQAAL 1307
 RESULT 14
 US-09-352-552-2
 ; Sequence 2, Application US/09352552
 ; Patent No. 6060264
 ; GENERAL INFORMATION:
 ; APPLICANT: Skatrud, Paul L.
 ; TITLE OF INVENTION: Multiple Drug Resistance Gene atlc of
 ; TITLE OF INVENTION: Aspergillus nidulans
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Eli Lilly and Company
 ; STREET: Lilly Corporate Center
 ; CITY: Indianapolis
 ; STATE: Indiana
 ; COUNTRY: U.S.
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/352,552
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/996,644
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Webster, Thomas D.
 ; REGISTRATION NUMBER: 39,872
 ; REFERENCE/DOCKET NUMBER: X-11765
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-276-3334
 ; TELEFAX: 317-276-2763
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1308 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-352-552-2
 Query Match 32.9%; Score 2208.5; DB 3; Length 1308;
 Best Local Similarity 40.3%; Pred. No. 3.1e-191;
 Matches 479; Conservative 223; Mismatches 456; Indels 31; Gaps 10;

```

OY 451 SLSMAGKTTALVGPSSGKSTVYGLVERFYMPKSTVLLDGHIDIDLNRKQOISLV 510
    11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 SLKVPAGKVTALVGGSGKSTVGLLEKMYNPTSGAIRIDGNLISELVNGMRNRVRLV 499
OY 511 SOEPVLFGTIYNKINHGILGTRENESEDKVELTENAKMANADFTALDEGETNY 570
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 500 QOEPVLFQSVVDNIRKGLVGTPTWENASREOMERVOEAKKLAYAEFFISELTDGIDTLL 559
OY 571 GORGFLSSGQKORAIARAVVSDPKILLDEATSDLTRKSEGVQOALERAEGRTTYV 630
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 560 GERGGILSSGQKORVAIARAVVSOQPKILLDEATSDLPDAETIYQKALDKAEGRTTYV 619
OY 631 IAHRLSTITAHRIYVYLVNGKIAEOGTDELDVDRGAKRYEAQR--INEOKEADALD 688
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 620 IAHKLATIKKADNIYVMSGHVTEOQTHESLIAKDGVYAGLVKIQMLAVNASAHNDVNEE 679
OY 689 ADAEDLTNDIAKIKTASASASSDLDGKPTTIDRTGTHKSVSSAILSKRPETPKYSLWT 748
    11 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 680 GGEDEVALLLEVTE-----TAVTRYPT--SIRGRANSIKDRDYEINIKHMD 722
OY 749 LUKFVASFNRR--PEIRPYMLIGLVSVLAGGO--PTQAVLYAKAISTLSLP--ESQYSKLKH 804
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 723 MAAALAYLVRECEPELKMAYLVYLLGGLGSCAMYPGQAILMSRYVEVFTLSGAMLDK--- 779
OY 805 DADFNLSMFEVVGIIQFTQSTNGAFAVCSERLIRRASTAFRTILRODIAFPDKEENS 864
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 780 -GDFYASMLIVLAAGCLICYLAVGATNTIAOHLSHMFRLILHMLRODIOFEDREENT 838
OY 865 TGAITSFELSTFKHLSGVSGLTITMTSTLGAIIITAIIGKMLAVCI--SVYVLL 923
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 839 TCAIVSRIDSYPHAILLELMGYNAIVYLVLOYVTCGILALFNMKGLVYVFGGIPPLV 898
OY 924 ACGETRYMLAOFQSRKSLAYEGSANFACCATSSIRTVASLTRERDWEIYHAQDQGR 983
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 839 GAGMAYRIRDSRLDROTSSKRYGTSSIASAVNAIRTVSSLAIEETVLRYTEELDHAVS 958
OY 984 TSLIVLRSSLLYXASQALVEFCVALGFWYGGTLGHHEYIDFRFFVCSSEILFGQSKG 1043
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 939 SSVKPMAMTICFGLTQCEIYFWQALGFYVGLVSLAGTSMVSEFVALSVFFAQASA 1018
OY 1044 TVFSEAPDGMKAKNAAEFRRLFEKRPQIDNMSSEGEKLETEGELEFENHFRYPTRE 1103
    1019 QLFQWSTSTTKINMTNYIAWLHQLOPYRETPENHDKPQSGAPLAMDVAFSTPLRD 1078
OY 1104 QPVLKGLDLYVPGQYVALVGPSSGKSTTALREFYDALIAGSILVYDKDISKINISY 1163
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1079 APLIGVNLKINKGQFIKAVGSSGCKSTMIAMLERFYDPTGSIITIDASTLIDIMPISY 1138
OY 1164 RSFLVLSOEPFLYOGTIRENILLGIVEDVPEEFLIKCKKCANITYDFINSLPEGEPTYV 1223
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1139 KRIYALVOOEPLTFQGTIDNLSLDGDAVKSDEQIESALRRANAMDFVSSLPOGLIYTPA 1198
OY 1224 GSKGMLSGGQKORVAIARALDRPKILLDEATSDALDESEKVVQALDAALARR--GRTT 1281
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1199 GSGGQSLSGGQKORAIARALRDRPKILLDEATSDALDESEKVIYQKALEGAARQGRDLT 1258
OY 1282 IAVARLSTIQKADVIYVFDQKIVESGTHSELVQKGRYELVLMQSL 1330
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1259 VAVARLSTIKDANYICVFEGGKIAEMGTHOELIYRGGLYRRMCEAQL 1307

```

```

ADDRESS: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,537
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X9212
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
TELEFAX: 317-276-1917
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-232-537-2

Query Match 29.6%; Score 1991.5; DB 1; Length 1302;
Best Local Similarity 35.3%; Pred. No. 1.7e-171;
Matches 471; Conservative 264; Mismatches 521; Indels 77; Gaps 23;

OY 43 LSPSSTATPADKEHRRPSSNNNAVSVN-EVDLIAHLHPD-----EROVIAKTL 93
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MSADGRRTVSSDYGH-PPSPMANEADHNEFTTINSQPSMWHTEFVFSKRIQOILL 59
OY 94 EIKVNIISFPGILMYRAT-KMDILINIVISTICAIIMAASTFORI--MLYQI--SYDEPYDEL 148
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 GKNPFKSYLDLFLKLVNDAKSKAVLMAIGLILMAAGCPPLIIGYIFQOITTSFPPEEDVL 119
OY 149 TKNVLYFVYLIGERYTVVSVGFIYTGEMATQKIREYLLSILRONIGYDKLAGAGV 208
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 RDRLQVLGVAGCYFIVTGYAIVMGITGEKISRRETLVRLGLLEQAYPD-IKDPDI 178
OY 209 TTRTADNLTIDOGISEKGLTLTLAFVTPAFITAYVYKWLALIGSSTIV--ALVIT 265
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 TMLTEKTEIADIGISEKGIYFQISISTYVAAFIVGFIIMAKLTGILFAPLMLALYV 238
OY 266 MGGSGOFLIKYSKSLDSYAGGTVAAEYVSSIRNATAFGTODKLAKOYEVHIDAEKMG 325
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 V--GSSRIAKYTKAATEYEAAGRIAESAIHAVKVVOAQGMENLSKEHYRLKLSARVA 296
OY 326 TKNQIVMGIMTAMGIMVSNGLGFWMKSRPLV-GAVDVCDILTVMALLIGSFSIGN 384
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 IRRSVSAEFMGLVYFTAYSANALAFWEGSRLAASGSNNACTYVAVVELLIDASFVVGQ 356
OY 385 VSPNQAETFNANAAAKIKFTIDR--QSPIDPYSNGK--TLDFHGHGIELRNVKTIYPSR 441
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 RGPFLGFRATAAAGESEYELLNHPQSELTINYSKGEQETESDMKADLFRNVTYVYPAR 416
OY 442 PEVTVMEDVSLSPAGKTTALVGPSSGKSTVYGLVERFYMPKSTVLLDGHIDIDLNR 501
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 417 TSARALEEMSLILKAGOMNAIVGTSGCKSTLVSLLRVYDSSQOLVITGSHDIDPWR 476
OY 502 WLRQOISLVSOQPVLPFTIYKINRGLIGTYENESDVKVELTENAKMANADFTTA 561
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 477 SLRKYTALVDDOSVLSGSEVLENISYGL---GEHSLSDVYLERCTEAKAANL-DFYDF 532
OY 562 LPEGYETNVGQGLF-LSGGQKORAIARAVVSDPKILLDEATSDALTRKSGVQOALE 620
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 533 LPQGIHTRIGNGYISLSGGQKORICLARALVKKPALLLDLDEPTALDANSBGLIMDAVK 592

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:58:28 ; Search time 57 Seconds
(without alignments)

2249.883 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721

Sequence: 1 MSPLETPNPLSPETAMREPAE.....YOKKGRYELVNLQSLGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3992	59.4	1307	2 T30882	multidrug resist
2	3068.5	45.7	1362	2 T41534	leptomycin B resis
3	2682.5	39.9	1408	2 T43261	multidrug resist
4	2493.5	37.1	1287	2 S55692	multidrug resist
5	2489	37.0	1280	1 DVH01	multidrug resist
6	2485	37.0	1276	1 DVH1C	multidrug resist
7	2471.5	36.8	1276	1 A34786	multidrug resist
8	2453.5	36.5	1276	1 DVH01	multidrug resist
9	2392.5	35.6	1277	1 JH0502	multidrug resist
10	2387	35.5	1276	1 DVH02	multidrug resist
11	2379.5	35.4	1279	1 DVH03	multidrug resist
12	2357	35.1	1278	2 S41646	P-glycoprotein - r
13	2357	35.1	1281	2 T48123	P-glycoprotein iso
14	2335.5	34.7	1104	1 DVH01A	multidrug resist
15	2232.5	33.2	1321	2 T42228	P-glycoprotein sis
16	2230.5	33.2	1321	2 T42842	probable ABC trans
17	2203.5	32.8	1278	2 E86155	probable ABC trans
18	2181.5	32.5	1286	2 T02187	multidrug resist
19	2173.5	32.3	1275	2 T31073	multidrug resist
20	2166.5	32.2	1292	2 T48007	P-glycoprotein hom
21	2146	31.9	1230	2 E85023	probable P-glycopr
22	2145	31.9	1289	2 D87789	protein C346.4 [1
23	2135	31.8	1294	2 T19982	hypothetical prote
24	2132	31.7	1229	2 F86155	probable ABC trans
25	2128	31.7	1229	2 D85023	P-glycoprotein-lik
26	2128	31.7	1229	2 T52319	P-glycoprotein-lik
27	2106.5	31.3	1323	2 H85202	hypothetical prote
28	2105	31.3	1302	2 S30327	multidrug resist
29	2098.5	31.2	1310	2 S30328	multidrug resist

30	2091.5	31.1	1283	2 A47377	multidrug resist
31	2079.5	30.9	1245	2 G86404	probable P-glycopr
32	2046.5	30.4	1327	2 T21268	hypothetical prote
33	2035	30.3	1247	2 F86405	probable P-glycopr
34	2024.5	30.1	1266	2 A42150	P-glycoprotein ppg
35	2010.5	29.9	1222	2 T14805	hypothetical prote
36	1998.5	29.7	1233	2 T04251	multidrug resist
37	1991.5	29.6	1302	2 T06165	multidrug resist
38	1989.5	29.6	1232	2 T06165	multidrug resist
39	1988.5	29.6	1321	2 T23476	hypothetical prote
40	1985.5	29.5	1321	2 S27337	multidrug resist
41	1971	29.3	1302	2 A41249	multidrug resist
42	1960	29.2	1318	2 T21266	hypothetical prote
43	1946.5	29.0	1316	2 B86240	protein F20B24.12
44	1934.5	28.8	1268	2 T22094	hypothetical prote
45	1933	28.8	1291	2 T21267	hypothetical prote

ALIGNMENTS

RESULT 1

multidrug resistance protein 1 - Aspergillus flavus

C:Species: Aspergillus flavus

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000

C:Accession: T30882

R:Robin, M.B.; Peery, R.B.; Skatrud, P.L.

Gene 200, 11-23, 1997

A:Title: Genes encoding multiple drug resistance-like proteins in Aspergillus fumigat

A:Reference number: 220917; MUID:98036972; PMID:9373135

A:Accession: T30882

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1307 <TOB>

A:Cross-references: EMBL:U62931; NID:92673946; PID:92673947; PIDN:AAB86655.1

C:Genetics:

A:Gene: mdrl

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match	Best Local Similarity	Score	DB 2:	Length
60	PKSSSSNNAVSVNEDALIAHLPEDEROVLTQLEIVNISFEGLMRATKMDLIVYI	59.48;	DB 2:	1307;
13	PKSPETGTTT-----GHSVSHAEEVLDRLHTPVSOIGFETVYRATKMDVALLFG	60.26;	Pred. No. 1.2e-23;	64
120	STICAIAA-----ASTFORIMLYQISDEYDELTKNVYFYVYIGIGFVT	60.26;	Pred. No. 1.2e-23;	64
65	SALAAAGGALPLETVLFGRLTSFPDLATHRTYDHFHHTLNVVYFYLGAAREYA	60.26;	Pred. No. 1.2e-23;	64
166	VYVTSVGTIVGEHAKQIREVYESTLRONIGYFDKLAGETVTRITADTNLIDGISE	60.26;	Pred. No. 1.2e-23;	64
125	YLAIVGFIYGDHVVQOIRVEYEQAILRONAFEDTLAGETVTRITADTNLIDGISE	60.26;	Pred. No. 1.2e-23;	64
226	KVGLTLTALFVFAFIYAVYKVMKALICSTIYALVLMGGSGQFIIRKSKSDSYG	60.26;	Pred. No. 1.2e-23;	64
185	KVGLTGLSTFVFAFIYAVYKVMKALICSTIYALVLMGGSGQFIIRKSKSDSYG	60.26;	Pred. No. 1.2e-23;	64
286	AGTVAEEVISIRNATAGTODKLAKOYEVLDEAEKMGTKNOIWMFGMIGAMGLWYS	60.26;	Pred. No. 1.2e-23;	64
245	RGASMAEDILDSIRIVAFNMOETLARKYESHLKABEGPMKSKYIFALMGALLCINYL	60.26;	Pred. No. 1.2e-23;	64
346	NYGLGFMMGSRPLVDGA--VDVGILTYLVALMILIGSFSLGNVSPNAQAFNVAANAATF	60.26;	Pred. No. 1.2e-23;	64
404	GTIRQSPLDYNSNGKTLDFEHEGIELRANKHYTPSPREYVYMDVSLSPAGCTTALV	60.26;	Pred. No. 1.2e-23;	64
365	GTIDQSPDLALSDGKTLFVGRNIVQNRHVPSPREYVYMDVSLSPAGCTTALV	60.26;	Pred. No. 1.2e-23;	64
464	GPSGSGKSTVGLVERFYMPVPGVTLDDHDKDLNRLMRLQOISLVQEPVLFCTTYK	60.26;	Pred. No. 1.2e-23;	64

Db 425 GSGSGSKSTIIILEEFLPVPAGTITMLDCHDIDQTLNIFRLROOMSIVSGEPRLFTTIAE 484
 QY 524 NTRHGLIGKYYENESDQKRELIENAKMANAHDFITLPEGEYFNWGRGLISGGQKQ 583
 Db 485 NTRVGLIGSREKESYETIRKVEEAAARNAHDFITLALPNQYDNI--ESFSLSGGQKQ 542
 QY 584 RAIAAAVSDPKITLLDEATGALDPTKSGVQAALERAEGRTTIVIAHRTSTIKTAN 643
 Db 543 RAIAARAIITKDKITLLDEATGALDPTKSGVQAALERAEGRTTIVIAHRTSTIKTAN 602
 QY 644 IYVLYNGKIAEGCTHDELVDRCGAKRKLYEQAORINDEKADALEDADELTN----- 696
 Db 603 IYVLYNGKIAEGCTHDELVDRCGAKRKLYEQAORINDEKADALEDADELTN----- 660
 QY 697 -----ADIAKITASSASDLDGKPTTIDTGTN-----KVSATLSLK-R 736
 Db 661 MFFPEKDYPGDESDIYST--LSDDASDI-----GHTGKQKPPVSRMSLSHLM 708
 QY 737 PEPTPKYSLMTLTKFVAFNRPETPYMLIGLVFVLAAGGQOPTQAVLYAKAISTLSPE 796
 Db 709 QPVKEEAYSFMTLTKFVAFNRPETPYMLIGLVFVLAAGGQOPTQAVLYAKAISTLSPE 768
 QY 797 SOYSLRHADAWSMFPPVVGITQITQSTGNAAVAVCEBRLIRARASTAFITLRODIA 856
 Db 769 LEPYLRHDAHWCMLEFVLMIGIVLVLSVQGTLEFVASEKKVYARSAQFAFVILHODIS 828
 QY 857 FPDKEENSTGALTSFSTETKHLGSGVGYTLGTLIMTSTTLGAALITAIAMKLTALYCI 916
 Db 829 FPDQENTGALTLATLSAGTKELTGISGVTGLTILVSVNLVYASIGVALYIGMKLTALYCI 888
 QY 917 SVYVPLALACGFREFYMLAOFQSRKLAIEGASNPACSEATSTIRYVASLTIREDDWEIYHA 976
 Db 889 SAVPALLMCGFRVYMLERFORAKKAYOESASACENASARTVVSILTMEALQSYOA 948
 QY 977 QLDAGRTSLISVLRSSLLIYASSQALVFECVALGFWMYGGTLLGHHEYDIFREYFQESLT 1036
 Db 949 QLRRLKSDILPITYVSSLLIYASSQALPFECVALGFWMYGGTLLGHHEYDIFREYFQESLT 1008
 QY 1037 FGAQAGVYFSAFAPDMGAKKNAFAERFLRDKPQIDNMSEGEKLETVGEIEERNAHF 1096
 Db 1009 FGAQAGVYFSAFAPDMGAKKNAFAERFLRDKPQIDNMSEGEKLETVGEIEERNAHF 1067
 QY 1097 RPTPEPQPVYLGDLIYVPGQYVALVNGSGGKSTTALLERFYDALAGSLIYVNGKDIS 1156
 Db 1068 RPTPEPQPVYLGDLIYVPGQYVALVNGSGGKSTTALLERFYDALAGSLIYVNGKDIS 1127
 QY 1157 KLINISYFSLVSGOEPTLYQGTIKENILIGIVEDDVEEELIYACADANIYDITMSP 1216
 Db 1128 TLEMSSYSHLALLISOEPTLYQGTIKENILIGIVEDDVEEELIYACADANIYDITMSP 1187
 QY 1217 EGFNFWVSGKGMISGGQKORAIARALIRDPKITLLDEATGALDSESEKVVQAALDAAA 1276
 Db 1188 EGFNFWVSGKGMISGGQKORAIARALIRDPKITLLDEATGALDSESEKVVQAALDAAA 1247
 QY 1277 RGRITIAVAHRLSTIOKADVIYVPOGKIVESGTHSELVOKKGRYIELVNLOS 1329
 Db 1248 RGRITIAVAHRLSTIOKADVIYVPOGKIVESGTHSELVOKKGRYIELVNLOS 1300

RESULT 2
 T41534
 Leptomycin B resistance protein, ABC transporter [Imported] - fission yeast (Schizosaccharomyces pombe)
 N:Alternate names: P-glycoprotein homology
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 Sequence revision 03-Dec-1999 #text-change 02-Feb-2001
 C:Accession: T41534, S20548
 R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: Z17318
 A:Accession: T41534
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-1362 <SEE>
 A:Cross-references: EMBL:AL031307; NID:93426126; PIDN:CAA20363.1; PID:93426129; GSPDB
 R:Experimental source: strain 972h; cosmid c663
 R:Nishitani, K.; Yoshida, M.; Nishimura, M.; Nishikawa, M.; Horinouchi, S.
 Mol. Microbiol. 6, 761-769, 1992
 A:Title: A leptomycin B resistance gene of Schizosaccharomyces pombe encodes a protein
 A:Reference number: S20548; MUID:92244046; PMID:1349418
 A:Accession: S20548
 A:Molecule type: DNA
 A:Residues: 1-1333, 'AC', 1336-1362 <NIS>
 A:Cross-references: EMBL:D10695; NID:9218549; PIDN:BA01537.1; PID:d1002012; PID:g218
 C:Genetics:
 A:Gene: SPDB:SPCC663.03; pm01
 A:Map position: 3
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; glycoprotein; membrane protein; nucleotide binding; P-loop
 F:438-641/Domain: ATP-binding cassette homology <AACLI>
 F:435-462/Region: nucleotide-binding motif A (P-loop)
 F:1137-1332/Domain: ATP-binding cassette homology <ABC2>
 F:1154-1161/Region: nucleotide-binding motif A (P-loop)

Query Match 45.7%; Score 3068.5; DB 2; Length 1362;
 Best Local Similarity 47.3%; Pred. No. 1.3e-177;
 Matches 656; Conservative 233; Mismatches 403; Indels 95; Gaps 18;

QY 14 AMREPAETSTEQOASTPHADDEKILSDLASPSTATPAPDKERHPPSSNNAAVANE 73
 Db 2 SLHKSSTSTYKND--NHSLD-----LSIKIPSENKFEETKE--NEASEHYVDYK 52
 QY 74 VDALIAHLPEDEROYLKQLEIKVYNISFF--GLMRYATKMDILIMVISTCAINA----- 127
 Db 53 -DPEQRYTP--EQELIYKQINDTPAKLSGYPRIISYADKMDIMQLAGTIGTIGAGLGP 110
 QY 128 -----ASTFORIMLQIISVDEYDELTKNVIYFVYLIGGEPTV 166
 Db 111 LMSLVSGQLAFAFDLMSAGKASPOH-----YVDH-----CLFTIYALGVGCS 157
 QY 167 YVSYVGFYITGEHATQKIREYLYLSIRQNIQNGYFDKLGAGEVTTITADTNLIDGISEK 226
 Db 158 YIYVTFIIAEERIAKRIKQDYLAHLSQNIQNGYFDKLGAGEVTTITADTNLIDGISEK 217
 QY 227 VGLTTLATATVTAFFIAYVYKWTALICSTIYALVITMGSGGQFITYKSKSLDSYGA 286
 Db 218 VGLTFPFIATVYSGFVLAIFIRMKFTLLISMPFPIGCGIGLGVPTFKNKQIAVAAE 277
 QY 287 GGYVAEEYISSIRNATAPGODKLAKOYEVLHDEAEKGTQKQIYVMEFMIGAMGLMYSN 346
 Db 278 SSTFEVEEYFNSIRNAPAGCTODILAKLNKTLITAAQRGINKALAMGLMGMFFVAYGV 337
 QY 347 YGLGFWMSRFLVQDANDVGGILYVLAAILIGSSSLGVSNAQAFITNANAANAAKITGTI 406
 Db 338 YGLFWFWGGRLLHGGDLDVSKLIGCFEAVLLASVLANISPKMOSFVSCASMAKIDDTI 397
 QY 407 DROSLPDYVSENEKTLDFEFGHIELRNKHYIPSPREYTVEDVLSMPAGKTALVGPS 466
 Db 398 DRVSPINAFITPGDYVNDIKKEIELKTRFYPYIPREYLVINDSLVCPGSKITALLGAS 457
 QY 467 GSGKSTVYGLVERFYPVYGTVLLDGHDKDLNDRWLNRQOISLYSQEPLYEGTITYKNIR 526
 Db 458 GSGKSTIIGLVERFYDPIGGQVFLDGDKDLRTLNVAISLRNQISLYSQEPLYEGTITYKNIR 517
 QY 527 HGLIGTKENESSEKVRRLIENAKMANAHDFITLPEGEYFNWGRGLISGGQKRIA 586
 Db 518 YGLPDTITGLTSKELERRYDAKLANAAYDFITLPEGEYFNWGRGLISGGQKRIA 577
 QY 587 IARAIVSDPKITLLDEATGALDPTKSGVQAALERAEGRTTIVIAHRTSTIKTANITYV 646
 Db 578 IARAIVSDPKITLLDEATGALDPTKSGVQAALERAEGRTTIVIAHRTSTIKTANITYV 637
 QY 647 LVNGKIAEGCTHDELVDRCGAKRKLYEQAORT-----NEQKADALEDADE-----D 693
 Db 638 LVNGKIAEGCTHDELVDRCGAKRKLYEQAORT-----NEQKADALEDADE-----D 697

QY 694 LTNADIKIT-----ASSASSDDGKPTTIDRTGTHKSVSAILSKRP-----E 739
 Db 698 DEDMDASLEAPMWSHNTDPTLNNKLNKEDKNVFEEDKTLQHVASEIIVPNLPADVGLN 757
 QY 740 TTPKYS-----LMTLLKPVASFNNP--EIPYMLGLVPSVLACGGOPQOAVLXA 786
 Db 758 EEPKSKSKSKNNHNSLTALMTFHSFVKMTITICLLIGLISMIGGAAPVQAAVFA 817
 QY 787 KAISTLSLPESQSKLRHDAFWSLAFVVGIIQFITQSTGCAFAVCSERLIRARASTA 846
 Db 818 RFLNIEF--DLSTDFLHKXNVFAVYWLILATVQFPAVATSNFMYAMEAVLQRIKYL 875
 QY 847 FRILLRODIAFDKENSTALSTFSLTERKHSVGSVGLTILMTSTLGAATIALA 906
 Db 876 FRLLRQDVFEFDESEVTAITSTSLQISLEGLSGPTLGFQTLNITLITSLIA 935
 QY 907 IGKMLAVCSVYPVLLACGFYRYMLAQOSRKLAVBESANPACERTSSIRVASTIR 966
 Db 936 TGMKGLVLTSTSVIITAGYVRALDVOEKLSAAVKESMAACESTSIRTVASLNR 995
 QY 967 ERDVMETIYHAQDAGRTSLISVRSLLYASSQALVFVAGLGFVGLGHHEYDIF 1026
 Db 996 EENVFAECDSILKPGRESALASLGSLEFSAAGVTFLLNLTFTWIGSTLMKGEYNIY 1055
 QY 1027 REFVCSSEILFGAQSAGTVESFADMGKANAAEFRRLLDRKQIDNWSEGEKLETV 1086
 Db 1056 QFTCFATVIFGIIQAGQFPGYSADYTKAKAAGEIKYLSBSKRPITWSTEGKVVSLQ 1115
 QY 1087 -GLEEFNNVHRYPTREQVYLRGLDITVPCGVVALVYSGCGSKSTTIALLEFYDAIA 1145
 Db 1116 SAAIEFQVSESYTRKHIVKRLNLTVRPGQVAVAGSSGCGSKSTTIGIEFYCDN 1175
 QY 1146 GSILVNDKDISKLNINSRFLSVSQEPFLYQGTIKENILIGIEDVPEEFILKACD 1205
 Db 1176 GAVLVQSVNVRDYNINRYQIALVYSEPTLYGTVARENITVLS-ASNDVSEEMIEACK 1234
 QY 1206 ANIYDFMSPDEGNNVYVSGKGLSGGQKORAVAIARALLDRPKILLDEATSAIDSE 1265
 Db 1235 ANIHEFTLPGNYNTLCGGKSSLSGGOKORAIARALLIRPKILLDEATSAIDSHSE 1294
 QY 1266 KVVQAAIDAAARGRTTAVAHRLSTIOKADIVYFEDQKIVESGTHSEIYOKKRYEYL 1325
 Db 1295 KVVQEAIDNAAASQRTTIAVIAHRUSSIODADCTIEVFGSGVIAEGTHAEVLAKQRYEYL 1354
 QY 1326 NLQSLGK 1332
 Db 1355 VEOGLNK 1361
 RESULT 3
 T43261
 multidrug resistance protein 1 - fungus (Filobasidium floriforme)
 C:Species: Filobasidium neoformans; Cryptococcus neoformans
 C:date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text-change 11-May-2000
 C:Accession: T43261
 R:Thornhill, S.J.; Peery, R.B.; Skatrud, P.L.
 Gene 201, 21-29, 1997
 A:title: Cloning and characterization of CneMDR1: a Cryptococcus neoformans gene encoding
 A:Accession: T43261
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1408 <THO>
 A:Cross-references: EMBL:U62929; NID:g2668552; PID:g2668553; PIDN:AA049889.1
 A:Experimental source: strain M1-106
 C:Genetics:
 A:Gene: mdr1
 A:Map position: IV
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; membrane protein; P-loop
 Query Match 39.9%; Score 2682.5; DB 2; Length 1408;
 Best Local Similarity 43.9%; Pred. No. 3; 5e-154;

Matches 597; Conservative 233; Mismatches 449; Indels 81; Gaps 16;
 QY 33 AADKKILSDLSAPSSPTA--TPADKREHPKSS--SSNNAVSVNEVDALIAHLPEDEROYL 89
 Db 69 AAEKISSSTIAASDILRNSPLEK---PISNAKSKSPIPKSKFDFLAKSKKEEBERK 125
 QY 90 KTDLEE--IKVNIISFGLWRYATKMDILLINVTICAIAAAS-----T 130
 Db 126 NKEKEKASVLPVPSFALPFAPILETIAVLGLVLAVAAGSQPLMTLIFGRILTSFT 185
 QY 131 FQRLMLQIIS-----DEFYDELKNVLYEYVIGIGEVYVYSTVGTFTY 175
 Db 186 NYAVIANQISQGLTPETSAALQAADLKTKQSGHMLLYLMAIIGMELATWLYMETINV 245
 QY 176 TGEHATOKIREYVLESTIRONIGYFDKAGEVTRITADNLNLODISEKVGILVTLA 235
 Db 246 TGLNSKRIRKRYTAAVLRQETLAFDDAGAEVATRIOTDHLVOEGSEKVALVFOYAG 305
 QY 236 TETVAFIIAYVYKWKALICSTIVALTVLVGGSGOFIIRKSKSLDSYGAGTVAEEVI 295
 Db 306 TFGVGFLVAFVRSPLRIGALVLSLPVIMLGGIMMTMAKGTALDIAKAGSLAEVYI 365
 QY 296 SSIRNATAFGTQDLAKOYEVHLDAEKWKTKNOIWMGFMIGAMPGLMYSVYGLGFWGCS 355
 Db 366 GSIRTVQAFGKEKILGDKFADHIEQSKIVGRKSGIFEGFGLSIMEFYIYAVALAFYVG 425
 QY 356 RFLVDGAVDVGDILVYLMALILGSPSLGNSVPMNAQAFNNAVAAAKITGCTIDROSPIDY 415
 Db 426 ILVSGAODSGIVINFEISLIGSSMMALPBLAAVYKAGAAKLPATIDRPAIDSA 485
 QY 416 SNEKRTIDHEFGHIELRNKHIIPSPREYVMEVDSLSMPAKTTALVPGSGSKSYVG 475
 Db 486 SEEGFKPGLGELISEFENVKFHPSPRPILPKGTTTFEAGKTFALVAGSGSGSTVYS 545
 QY 476 LYERFAPVRKTVLLDGDIDOLNRLRQOISLVSQEPFLFGTTIKINAGLIGTYE 535
 Db 546 LIERFYDPSGVVLLDGDIDLSLMLMDRQIGVSOEPTLFGTVYRNVHGLIGTYE 605
 QY 536 NESDEKVELTENAKMANADFTALPEGETVNVGGFLLSGGOKORAIARAVASDP 595
 Db 606 NASLEKFEVLYKACVDANANFIMKLPQGDITVNGEGMILSGGOKORAVAIARAVSDP 665
 QY 596 KILLDEATSAIDTKSSEGVQALERAEGRTTIVIAHRLSTIRTAHNIYLVNGKIAEQ 655
 Db 666 RILLDEATSAIDTKSSEGVQALDALKASRGRTTITIAHRLSTIRADRIYVNGGEVLEQ 725
 QY 656 GTHDELV-DRGGAIRKLYEADRIEOKRADLE-DADAEDLTNADIAKITAASASDLD 713
 Db 726 GSHNDLLANENGPYAOALVNNOKLQEAALAEALQYDDIEDPDA-----VFIGSSSPMQ 779
 QY 714 GKPTTIDRTGTHKSVSSA-----ILSKRPETT-----PKYSLMTIL-----KVFASF 756
 Db 780 EKDQOLHRAVYGRSLAISTAMDIOAKRAEYVAGEDKIPSSGFLARLRMSADKFI--- 836
 QY 757 NRPEIPYMLGLVPSVLACGGOPQOAVLYAKAISTLSLPESQSKLRHDAFWSLAFVYV 816
 Db 837 -----YIAFIAICAGMYPSLILFGKALDSEFIQDP--AEIRALRSRLMWYIT 887
 QY 817 GIIDFIQSTGCAFAVCSERLIRARSTARTILRQDIAFDKENSTALSTFSLTSET 876
 Db 888 ALAAAFVIFQSAFESRAGWDLNGLVRKKLTATLRHDIEMFDERNSTGAVTSLADOP 947
 QY 877 KHLGSGVGLTILMTSTLGAATIALAIGWKIALVCISVNVPLACGYYRYMYLAQF 936
 Db 948 QKVGGLGPTLIGYVQSGATIGGCTIGLCGPLALIGLICIFILVSGGTYIRLKVYVLK 1007
 QY 937 QSRKLAIVEGSAFACERTSSIRTVASTLRERDWEIYHAQDAGRTSLISVRSLLY 996
 Db 1008 DQRRKILHAASAHALASEAAGVKTVAALTRBKOVRIYSEALKAPMKLNFTSIKSOGLF 1067
 QY 997 ASSQALVFECVAGLGFVGGTILGHHEYDIFREFVCSSEILFGAQSAGTVESFADMGKAK 1056
 Db 1068 AASGGLTFCTIALVFIYIGALWITDAKYSTASFTVLSIVPASIOAGVTFVVDASKAN 1127

A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
 A:Reference number: A34914; MOID:90094448; PMID:1967175
 A:Accession: A34914
 A:Molecule type: DNA
 A:Residues: 1-1280 <CH2>
 A:Cross-references: GB:M29447; GB:J05168; NID:9187496; PIDN:AAA59576.1; PID:9386862
 R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.
 submitted to JIPID, April 1991
 A:Reference number: PS0162
 A:Accession: PS0162
 A:Molecule type: DNA
 A:Residues: 1-22 <KTO>
 R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.
 submitted to the EMBL Data Library, April 1991
 A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) expression
 A:Reference number: S15500
 A:Accession: S15500
 A:Molecule type: DNA
 A:Residues: 1-22, 'R' <R12>
 A:Cross-references: EMBL:X58723; NID:934522; PIDN:CAA1558.1; PID:934523
 R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I.
 Cell 47, 381-389, 1986
 A:Title: Internal duplication and homology with bacterial transport proteins in the mdrl
 A:Reference number: A25059; MOID:87028230; PMID:2876781
 A:Accession: A25059
 A:Molecule type: mRNA
 A:Residues: 1-184, 'Y', '186-1280 <CH2>
 A:Cross-references: GB:M14758; NID:9187468; PIDN:AAA59575.1; PID:9307180
 R:Chambers, T.C.; Pohl, J.; Glaser, D.B.; Kuo, J.F.
 Biochem. J. 299, 309-315, 1994
 A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase of
 A:Reference number: S43838; MOID:94220047; PMID:7909431
 A:Accession: S43838
 A:Molecule type: protein
 A:Residues: 656-689 <CH2>
 R:Geckeler, V.; Weger, S.; Probst, H.
 Biochem. Biophys. Res. Commun. 169, 796-802, 1990
 A:Title: mdrl/P-glycoprotein gene segments analyzed from various human leukemic cell lines
 A:Reference number: I52238; MOID:90290529; PMID:1972623
 A:Accession: I52238
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 178-215 <RBS>
 A:Cross-references: GB:M37724; NID:9183537; PIDN:AAA88047.1; PID:9553314
 A:Accession: I65204
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 800-856 <RBS>
 A:Cross-references: GB:M37725; NID:9183538; PIDN:AAA88048.1; PID:9553315
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
 structurally and functionally unrelated lipophilic antitumor drugs.
 C:Genetics:
 A:Gene: GDB:PGY1; MDRI
 A:Cross-references: GDB:120712; OMIM:171050
 A:Map position: 7q21-7q21
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein;
 F:1-638, 653-1280/Region: duplication
 F:49-350/Domain: hydrophobic <HB1>
 F:351-637/Domain: hydrophobic <HB1>
 F:410-604/Domain: ATP-binding cassette homology <ABC1>
 F:427-434/Region: nucleotide-binding motif A (P-loop)
 F:551-555/Region: nucleotide-binding motif B
 F:638-708/Domain: linker <LIN>
 F:709-993/Domain: hydrophobic <HB2>
 F:994-1280/Domain: hydrophobic <HB2>
 F:1053-1249/Domain: ATP-binding cassette homology <ABC2>
 F:1070-1077/Region: nucleotide-binding motif A (P-loop)
 F:1196-1200/Region: nucleotide-binding motif B
 F:91, 94, 99/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:433/Binding site: ATP (Lys) #status predicted
 F:661, 667, 671/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status exp
 F:667, 671, 683/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status

F:1076/Binding site: ATP (Lys) #status predicted
 Query Match 37.0% Score 2489; DB 1; Length 1280;
 Best Local Similarity 42.1%; Pred. No. 1.6e-12;
 Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;
 QY 94 EEIVNISFFGLMRYATKMDILINWISTICAIAASTFORIMLY----- 137
 DB 28 KKKPTVSVFEMFRYSNNLDKLYVAVGTALIIHGAGLPLMLVFGKMTDIFANANGIED 87
 QY 138 -----QISYDEYF---DELTKNLYFVYVIGIEFTVVSYVGEITYTGEHAKRI 184
 DB 88 LMSNITNRKNDINDYGFENLNEEDMTRAYYSGAGVLAIVAOVSFWCLAGAQIHKI 147
 QY 185 REYVLESTIRONICGYEPDKLGGEVYTRTADPNLQGISKVGITLALFTYATPIA 244
 DB 148 RKQFFHAIKROEIGMPVDHVDGELNTRTDVSKINEISGKIDGMFQSMATFTFTGVG 207
 QY 245 YVKYWKALICSSYTVAVLITMGSGSPFIIRYKSKSDSYGAGTVAAEEVTSIRNATAF 304
 DB 208 FTRGKTLVILVLAISPVIGLSAAVWAKLISFTDKELLAYAKAGAVAEVLAIRTYAF 267
 QY 305 GTQDLAKQYEVHLDKAEKMGTKNOIYWGEMIGAMFGIMYSNYGLGFWMGSRFTVDAVD 364
 DB 268 GGQKKELEERYNKNEEAKRIGIKKAITANISGAFLITVASYALAFWYGTTLVSGEYS 327
 QY 365 VGDILTVMLALISFSGSNPNAQFTNVAATAKFGIIDQSPIDPISNEGKTIIDH 424
 DB 328 IGQVTVFVSVLIGAFSGQASPSIEAFANRGAAYELFKIIDKPSIDSYSKSGHKRDN 387
 QY 425 FECHIELNRYNHIYSPREYVMDVSLSPAGTKTALVPGSGSGKSTVGLVERFVYV 484
 DB 388 IKGNLEFRNRYNFIYSPSKREVILKGNLKNYSGQTVLVGSGGKSTVGLMQRIDPT 447
 QY 485 RGYVLGDHDIKDILNKLKQQLSVQSEPVLFSTTYIKNRHGLIGTRYNESEDKYRE 544
 DB 448 EGMVSVGDQIRITINVRFLREIIGVSOEPLFTTAENIRYG----RENVWDE---- 499
 QY 545 LIENAAKMANHDTLPLPEYENNVGQGLSGGQKQRIARAVYSDKTIILDEAT 604
 DB 500 -IEKAVENANNYDFIMKIPKFDVLVGERGQSLSGGQRIARALVRNKKIILLDEAT 558
 QY 605 SALDTKSGVQVQALERAEGRTTVIAHRLSTIKTAHNIYVAVNGIAEGTDELVD 664
 DB 559 SALDTESAVVQVADARAKRRTIIVIAHRLSTIRNMDVIAGFPDGVYVEGNDELKKE 618
 QY 665 GGAVRKLVQAEORINEQKEDALADADADLTNADIATIKTASASSSDLDGKPTTIDRTGT 724
 DB 619 KGIYFKLVMTQAGNEVE--LENA-----ADESKSEIDALEMSSNDSRLRKST 668
 QY 725 HKSVSAIISKR-----PETPKYSIMTLLFVASFNPPELPIYMLIGLVFVLAGGG 777
 DB 669 RRSVRSQADRKISTEALDESIPVSEWRIMK---LNTLTPYFVGVGCAIIMGGL 724
 QY 778 QPTQVAVYAKAISTLSPESQSKLRHADFWMSLMFEVYGIPTFTOSTNAAAVCSER 837
 DB 725 QPAVAIFSKITIGVTRIDPEK-RQNSNLSFLFALAGISITFTPLFGFTGKAGEI 783
 QY 838 LIRRASTAFETILRODIAFPDKENSTGALTSLETYKHLGVSQVYLGITIMTSTTL 897
 DB 784 LTKRLRYMVFERSMLROVSWPDDPKTGALTTRLANDAAQVKAIGSRLAVITONANL 843
 QY 898 GAATITLALGKMLAVICISVVPVILACGFRFIMLAQFQSRSLANEGSNPCEATSS 957
 DB 844 GTGIIISIFYGMQLTLTLLAVPIITLAVGVEMKLSGQALKDKKEGAKAKITAEIEN 903
 QY 958 IRTVASTIRERDWEIYHAQDAGRTSLISVLRSSLLIYASSQALVEFCVALGFYGGTL 1017
 DB 904 FRTVYSLTQEKFEHMTAOSLOVYRNSLRKAHIFGITFTSQAMMFTSYAGCRFGAYL 963
 QY 1018 LGHHEYDIFREYFCSEILFGAQSAGTVFSFADPMGKAKNAAEFRRLFDKRPDIOWMSE 1077
 DB 964 VAHKIMSFEDVLIVFSAVVFAGMAVGVSPADPYAKAKISAHIIMTIETKPLIDYST 1023

RESULT 7
multidrug resistance protein 1a - mouse
A:34786
C:Species: Mus musculus (house mouse)
C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 02-Feb-2001
C:Accession: A34786; A35671
R:Devault, A.; Gros, P.
MOL. CELL. BIOL. 10, 1652-1663, 1990
A:Title: Two members of the mouse mdr gene family confer multidrug resistance with overlapping
A:Reference number: A34786; MUID:90205845; PMID:1969610
A:Accession: A34786
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1276 <DEV>
A:Cross-references: GB:X30697; NID:g19911; PIDN:AAA39517.1; PID:g387429
R:Han, S.I.H.; Cohen, D.; Kirschner, L.S.; Lothstein, L.; Hattstein, M.; Horwitz, S.B.
MOL. CELL. BIOL. 10, 3596-3606, 1990
A:Title: Structural analysis of the mouse mdr1a (P-glycoprotein) promoter reveals the ba
A:Reference number: A35671, MUID:90287150; PMID:1972547
A:Accession: A35671
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-525, 'HV', 528-938, 'S', 940-1035, 'V', 1037-1276 <HSU>
A:Cross-references: GB:X33561; NID:g199104; PIDN:AAA39514.1; PID:g387427
C:Superfamily: multidrug resistance protein, ATP-binding cassette homology <ABCl>
C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-loop
F:406-600/Domain: ATP-binding cassette homology motif A (P-loop)
F:423-431/Region: nucleotide-binding motif B
F:547-551/Region: nucleotide-binding motif A (P-loop)
F:1049-1245/Domain: ATP-binding cassette homology motif A (P-loop)
F:1066-1074/Region: nucleotide-binding motif A (P-loop)
F:1192-1196/Region: nucleotide-binding motif B
F:429/Binding site: ATP (Lys) #status predicted
F:1072/Binding site: ATP (Lys) #status predicted

[illegible]

Db 1023 KPNMEGNVQSGFVFNPTFRSIPVLOGSLSEVKKGQTLALVSSGCKSTVOLLERF 1082
 QY 1141 YDAAGSILVDGKDISKLNINSRSLVSOEPTLYOGTKEINLLGIVEDVPEEPLI 1200
 Db 1083 YDPMAGSVFLDGKEIKOLNVMQLRAOLGIVSOEPLIFDQSIENATVAGNSVSVSEELV 1142
 QY 1201 KACRANAIYDFMSLPEGENIVYVSGKGMLSGGOKORVAIAALLRDKPILLIDEATSL 1260
 Db 1143 RAAKAAHITQFIDSLPDKRNTVCGDKGTQLSGGOKORIAIAALVQPHILLIDEATSL 1202
 QY 1261 DSESEKVVQALDAARGRRTTAVAHRLSTIQKADIVYFVDGKIVESGTHSELVQKKR 1320
 Db 1203 DTESKVVQOALDKRAREGRTCTVIAHRLSTIQKADIVYFVDGKIVESGTHSELVQKKR 1262
 QY 1321 YVEIVNQLSLGR 1332
 Db 1263 YFSMVSVQAGAK 1274
 RESULT 8
 DVM51
 multilidrug resistance protein 1 - mouse
 N:Alternate names: P-glycoprotein 1
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
 C:Accession: A33719; A25057; I57510
 R:Raymond, M.; Gros, P.
 Proc. Natl. Acad. Sci. U.S.A. 86, 6488-6492, 1989
 A:Title: Mammalian multilidrug-resistance gene: correlation of exon organization with stru
 A:Reference number: A33719; MUID:89367274; PMID:2570420
 A:Accession: A33719
 A:Molecule type: DNA
 A:Residues: 1-1276 <RAY>
 R:Gros, P.; Croop, J.; Housman, D.
 Cell 47, 371-380, 1986
 A:Title: Mammalian multilidrug resistance gene: complete cDNA sequence indicates strong ho
 A:Reference number: A25057; MUID:87028229; PMID:3768958
 A:Accession: A25057
 A:Molecule type: mRNA
 A:Residues: 1-1276 <GRO>
 A:Cross-references: GB:M14757; NID:g199100; PIDN:AAA39513.1; PID:g387426
 R:Raymond, M.; Gros, P.
 Mol. Cell. Biol. 10, 6036-6040, 1990
 A:Title: Cell-specific activity of cis-acting regulatory elements in the promoter of the
 A:Reference number: I57510; MUID:91042535; PMID:2248681
 A:Accession: I57510
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-21 <RES>
 A:Cross-references: GB:M60348; NID:g199102; PIDN:AAA39513.1; PID:g554199
 C:Comment: This is an integral membrane protein overproduced in multilidrug-resistant cell
 C:Genetics: structurally and functionally unrelated lipophilic antitumor drugs.
 A:Gene: mdrl (pgp1)
 A:Introns: 21/2; 38/3; 92/1; 111/2; 175/2; 233/3; 274/2; 332/3; 370/3; 407/3; 449/3; 517
 C:Superfamily: multilidrug resistance protein; ATP-binding cassette homology
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr
 F:1-637; 653-1276/Region: duplication
 F:409-603/Domain: ATP-binding cassette homology <ABC1>
 F:426-433/Region: nucleotide-binding motif A (P-loop)
 F:550-554/Region: nucleotide-binding motif B
 F:1051-1247/Domain: ATP-binding cassette homology <ABC2>
 F:1068-1075/Region: nucleotide-binding motif A (P-loop)
 F:1194-1198/Region: nucleotide-binding motif B
 F:73-91; 96-103/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:433/Binding site: ATP (Lys) #status predicted
 F:1074/Binding site: ATP (Lys) #status predicted
 Query Match 36.5%; Score 2453.5; DB 1; Length 1276;
 Best Local Similarity 40.9%; Pred. No. 2.3e-140;
 Matches 520; Conservative 232; Mismatches 461; Indels 57; Gaps 7;
 QY 90 KTOLEIVNIVISFGIMKRYATKMDILINIVISTICAIAAASFORIML----- 136

Db 23 KREKREKPAVGVGFGRFADWLDKICMLTGTLAAIHGTLPLMLTVGNMTDSFTKAE 82
 QY 137 -----YOISYDEFDLCKNVLYFYLIGFEMVYVSGFYTGEHA 180
 Db 83 ASILSITNOSPNSITLISNSLSLEEMAIYAYTIGAGVLAIVYISVSLCAAGHQ 142
 QY 181 TOKIREYLESILRONIGYFDKLAGEVYTRITADTNLIDGISEKVGTLTALATFYTA 240
 Db 143 IKRIKQFPFHALIMNDEIGFVDHOGELNTRLDVSKINDIGIKGFQSIPTPLAG 202
 QY 241 FIITAYKYKMLLICSSTVALVLMGSGSQFTIKSKSLDSYGAGTVAEEVSSITN 300
 Db 203 FIIGFSGKRLVTLVLAASPLGLISALMAKVLSTNNLEQAYAKAGVAEEVLAAT 262
 QY 301 ATAFGTQDKLAKQYEHVHDEAEKMGTKNOIVMGFMIGAMEGLMYSNYGFGMGSREFLD 360
 Db 263 VIAFGQCKELEERKNKNELEAKNVGIKRATISISIGIVLYVASYALAFYGTSLYS 322
 QY 361 GAVDVQDILITVLAAILIGSFLGNVSPNAQFTNVAATAKIFGTIDRQSPDPYSNECK 420
 Db 323 NEYSIGEVLTVFESILGTFSIGHLAPNIEAFANRGAFAELFKILLDNPSIDSTKGY 382
 QY 421 TLDHEGHELMKNKHIYPSREYVMEVDSLSMAGKTTALVPGSGSKSTVGLVERF 480
 Db 383 KPDISGMNLEFNKVNHPNPSRSEVOILKGLNKKVSGQTVLVAGNSGCKSTVQLMORT 442
 QY 481 YMPKRGTVLLDDHDKIDLRLWLROQISLVSOEPLVFGTITKYNRHLGIGTKYENESD 540
 Db 443 YPDEGVYSIDQDRTITVRLRELIGVSOEPLFATTIENIRYG-----RED 493
 QY 541 KYRELLENAAKANAHDFTALPEGETNVGQGFLLSGGOKORIAIARAVSDPKILL 600
 Db 494 YTMDEIEKAVKANAADFIMKLPDQFLVGERGAQLSGGOKORIAIARAVRNPRIKL 553
 QY 601 DATATADPKKSGVQAALERRAEGRTTVIAHRLSTITAINIYLVNKGIAEGTDE 660
 Db 554 DEATSLDTESEAVVQALDKAREGRTTVIAHRLSTVANAIVAGFDGIVTEGNDHE 613
 QY 661 LVDRGAVYKLEVAQ-RINEQKAD---ALPDADIEDTNDIARIKIKTASASDLDCK 715
 Db 614 LIREKGIYFKLMTQTRGEIERPKNNAVGSODITASELTSSEKSPILRRIYYSVHK 673
 QY 716 PPTIDRTGTHKSVSSAIIKRPETTPKYSMTLLKTVASFNRPETPYVLIGLVSVLAG 775
 Db 674 QDOERRLSKAEVAD-----EDVPLVSFWRIL---NLMSWPYLVGLVAIVING 720
 QY 776 GGQPGVAVLYAKAISTLSPEQYSKLRHDADFMGLMFPVGLIQTGSTGAFVACS 835
 Db 721 CIQPVFAIVFSRIVGVFSRDDDEHK-RONCNLFSLFELVMGLISFYVFFQGFTEGKAG 779
 QY 836 ERLIRARSTARFTLLRODIAFFDKEENSTGALSTFSLTETHLSGVSGVTGLTLMST 895
 Db 780 ELTRKVRNRVKSMLRODISWFDHKNSTGSLTRIASDASVYGAMARLAIVTQYNA 839
 QY 896 TLGAALITLAIKMLALVCISVPVLLACGFYRYTMAQFQSRKLAEGSANFACENT 955
 Db 840 NGGTGIVSLVYGQDITLLVVIIPLIYVIGITTEKMLISQALKKKOLEISGKIATFAT 899
 QY 956 SSIRVYASTRRERDWEIYHQAOLDAGRSLISVLRSSILVYASQALVFCALAFWKG 1015
 Db 900 ENKRTIVSLTRQKQKETMAQSLQVPRNAMKKAHVGTTFSTQAMMFSTAAQFRFKA 959
 QY 1016 TLGHEHYDIFFPFCSEILFGAOSAGTVFSFADPMCKAKNAAEFRILPRKQIOWM 1075
 Db 960 YVLAQQLMFENVMVLVSAVVGAAAGNTSSFADYAKAKASASIIIRIIEKTEIDSY 1019
 QY 1076 SEEGKLETVGEIEFRNVHFRYPTRPEQPVLRGLDLYVQGYVALVPGSGCKSTYTA 1135
 Db 1020 SNEGKPTLLEGNVKPNGVQFVYPRPNIPVLQGISLEVKKQTLALVSSGCKSTVQ 1079
 QY 1136 TLERVDIAAGSILVDGKDISKLNINSRSLVSOEPTLYOGTKEINLLGIVEDVPE 1195

Db 1080 LLERFDPNAGSFVLDGKETKOLNVOMLRAHLGIVSOEPIFLDPCSTAEINAYGDNSRAVS 1139
 QY 1196 EEFLLKACKDANIYDFIMSLPEGFNTVVGSGKGMGOKORVAIARALLRDPKILLDE 1255
 Db 1140 HEEIVAAAEANIHQFIDSLPKYNTNRVGDKGQLSGGOKRIARALVRQHIILLDE 1199
 QY 1256 ATSDSESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVFDQGTVESGTHSELV 1315
 Db 1200 ATSDTESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVFDQGTVESGTHSELV 1259
 QY 1316 OKRGYRELV 1325
 Db 1260 AOKGITYSMV 1269

RESULT 9

JH0502

P-glycoprotein - rat

M:Alternate names: multidrug resistance protein mdr1b

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 02-Feb-2001

A:Accession: JH0502; S22353; S22352

R:Sllyerman, J.A.; Raunio, H.; Gant, T.W.; Thorgerlsson, S.S.

Gene 106, 229-236, 1991

A:Title: Cloning and characterization of a member of the rat multidrug resistance (mdr)

A:Reference number: JH0502; MUID:92039081; PMID:1682220

A:Accession: JH0502

A:Molecule type: mRNA

A:Residues: 1-1277 <SII>

A:Cross-References: GB:M62425

R:Deuchars, K.L.; Duthie, M.; Ling, V.

Biochim. Biophys. Acta 1130, 157-165, 1992

A:Title: Identification of distinct P-glycoprotein gene sequences in rat.

A:Reference number: S22351; MUID:92223089; PMID:1348630

A:Accession: S22353

A:Molecule type: DNA

A:Residues: 1212-1226, 'T', 1228-1277 <DEU>

A:Cross-References: EMBL:X61104; NID:956890; PIDN:CAA43416.1; PID:91334219

A:Accession: S22352

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1212-1226, 'T', 1228-1270, 'SV', 1271-1277 <DE2>

A:Cross-References: EMBL:X61103; NID:956888; PIDN:CAA43415.1; PID:91334218

A:Comment: This protein has the nucleotide binding motifs and ATP binding active transp

C:Genetics:

A:Gene: mdr1b

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-loop

F:409-603/Domain: ATP-binding cassette homology <ABC1>

F:426-434/Region: nucleotide-binding motif A (P-loop)

F:550-554/Region: nucleotide-binding motif B

F:1053-1248/Domain: ATP-binding cassette homology <ABC2>

F:1070-1078/Region: ATP-binding cassette homology <ABC2>

F:1195-1199/Region: nucleotide-binding motif A (P-loop)

F:432/Binding site: ATP (Lys) #status predicted

F:1076/Binding site: ATP (Lys) #status predicted

Query Match 35.6%; Score 2392.5; DB 2: Length 1277;

Best Local Similarity 40.4%; Pred. No. 1.2e-136;

Matches 513; Conservative 236; Mismatches 459; Indels 61; Gaps 10;

QY 94 EEIKVNSFGLMRYATKMDILIMVISTICAIAASTFORIML----- 136
 Db 26 KEKPPAVGIGFMRVADMDKLCMALGTAAIHGTLPLMLLVFGVMTDSFPPSRPHS 85
 QY 137 -----YQSYDEYDELTKNLYEYLLGIGEVTVYVSVGIIYGEHAFOK 183
 Db 86 DRAITNOSSEINSTRVSDTSLIEDDMATAYYIGAGVIVAYIQVSLCLAAGROIKH 145
 QY 184 IREYVLESTLRONIGFVFDKAGAEVTRITADINLQDGISEKVGILLTLATPVAFTI 243
 Db 146 IROKFFHAINMOELGMDVNDAGELNRLTDVSKINDIGDKLGMFFQSTITFSAGFI 205

QY 244 AYVKWKLALICSTIYALVLMGGGQFIKKYSKLSDTYGAGGYAAEYVISTIRNATA 303
 Db 206 GFTSGWMLTVIILVAVSPILICSSAMAKVLTSPFNKELQAVAAEVEALAIRVIA 265
 QY 304 FGNODKLAKOYEHVLDAAEKWGTFRNOIVMGFMIGMGLVANSYNGLGFWMSRRLVYGA 363
 Db 266 FGGQKLEKRYNKNLEKAKRNGIKALTANISIGIALLVYASTALFMWGTSLVLSNEY 325
 QY 364 DVGDILVLAAILIGSFSLGNVSPNAQFTNAAVAAAKIGTIDROSLDPYSNEGTLD 423
 Db 326 SIGQVLTFEFSILLGTSTIGHLANPTEAFANARAALAEIKIIDNESIDISFKGKHPD 385
 QY 424 HEEGHIELRVNKHITPSRPEVTYVEDVSLSPACKTTALVGPSSGKSTYVGLVERFYP 483
 Db 386 SIMGNLEFKNVYFNPSPRSEVKILKGLNLKVKSGQVAVLVGNSGCKSTVOLLQRLYDP 445
 QY 484 VRGTVLLDGDIDKDLNRMVROQISLVSOEPLVFGTITYNIRGLGTENSEGEKVR 543
 Db 446 IEGEVSIDGDIRTNRYREITIGVVSQEPVLEATTIENIRYQ-----RENTYKDE-- 498
 QY 544 ELIENAKMANANDFITALPEGYETNNGORGLSGGOKRIARAVVSDPKILLDEA 603
 Db 499 --IEKAVKENANVDFIMKLPKFTLVGERQAOLSGGOKRIARALVRPKILLDEA 556
 QY 604 TSALDTSEGVQALERAARGRTTIVIAHRLSTIKTAHNTVLVNGKIAEGTHDELVD 663
 Db 557 TSALDTSEAVVQALDAAARGRTTIVIAHRLSTIVRNADVAGPDGVIYVQGNHEELM 616
 QY 664 RCGAYRKLVERQ-RINE-----QKADALEDAADLDLNNADAKIKTASASASDLDGKPTT 718
 Db 617 EKGIFKLVMTQTGNIEEPGNNAYESQSDGASLSEESKPLIRRSIRSRHRQDQ 676
 QY 719 IDRTGKRSVSATLSKRPPETPKYSLMTLTKFVASPNREIYMYLIGLVSVIAGGQ 778
 Db 677 ERLISKEDVD-----EDVPMVSFWQILK-----LNISEMYLVVGVCAIYNGCIO 723
 QY 779 PTOAVLYAKAISTLSESOYSKLRHDAFWSIMFVVGIIQFTOSTNGAFAVCSERL 838
 Db 724 PVFAIVSKIVGVFSRDDDEHTQR-NONLESLFLVWGMISFYTFYFQGTFFKAGAIL 782
 QY 839 IRRARSTAFRIILQODIAFPDKENSSTGALTSFISTEKHLISGVYTLGITIMTSTLG 898
 Db 783 TKRLRYVFKSMQDLSWEDDHNTGSLTTRLASASANKGAMGSLAVYTONVANLG 842
 QY 899 AAILIALAI--GKRLAVCSIVPVLLACGYRPRYMLAQFOSRSKLAVEGSANFACETS 956
 Db 843 TGIILSLVLYGMOLLTLVITLPIVLVGIIEKLLSGQALKKKELEISGKIAIEAIE 902
 QY 957 SIRTVASLTRERDWEIYHQAQGRSLISVLRSSILVYASQALVEFCVAGFWYGGT 1016
 Db 903 NFRVVSILTRQKRETVYAQSLQIPYRNALKKAHVFGITFAQTQAMTYFVSAACFRCAY 962
 QY 1017 ILGHEVDIFRFFCFSEILFGAQSAGVPSFAADMKAKMAADEFRLPRKQIDNWS 1076
 Db 963 LVABELMTFENVMVLFSAVVFAGMAAGTSSFADYAAKVASASHIIGILEIKEDISYS 1022
 QY 1077 EEEGKLETFEGEIEFRVWVFYPTRPQPVLRGLDLYKPPQYVALVGPSSGKSTIATL 1136
 Db 1023 TEGKLPKMLBENAVFNGKVFYPTRPNIPLYQGLSFYKKQOTLRVLSGSGKSTVQVL 1082
 QY 1137 LERFYDAIAGSILVDGDISKLINSYRSLSLVSOEPLVYQGTIKENILGIVEDDVE 1196
 Db 1083 LERFYDPAAGVFLDQKEIKQLNVQVRA-LGVISOEPIFLDPCSTAEINAYGDNSRVVSH 1141
 QY 1197 EFLKACKDANIYDFIMSLPEGFNTVVGSGKGMGOKORVAIARALLRDPKILLDEA 1256
 Db 1142 EEIVRAAREANIHQFIDSLPKYNTNRVGDKGQLSGGOKRIARALVRQPHILLDEA 1201
 QY 1257 TSALDSESEKVVQALDAAARGRTTIAVAHRLSTIQKADVIYVFDQGTVESGTHSELV 1316
 Db 1202 TSALDTESEKVVQALDAAARGRTTIVIAHRLSTIQKADVIYVFDQGTVESGTHSELV 1261
 QY 1317 KKGRYRELV 1325

Db 1262 QKGIYFSMV 1270

RESULT 10

DVMS2
multidrug resistance protein 2 - mouse
N:Alternate names: P-glycoprotein MDR2
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: A30409; S70711
R:Gros, P.; Raymond, M.; Bell, J.; Hausman, D.
Mol. Cell. Biol. 8, 2770-2778, 1988
A:Title: Cloning and characterization of a second member of the mouse mdr gene family.
A:Reference number: A30409; MUID:88302195; PMID:3405218
A:Accession: A30409
A:Molecule type: mRNA
A:Residues: 1-1276 <HSU>
A:Cross-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428
R:Kirschner, L.S.
Nucleic Acids Res. 24, 2829-2834, 1996
A:Title: De novo generation of simple sequence during gene amplification.
A:Reference number: S70711; MUID:96313253; PMID:8759018
A:Accession: S70711
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 43-92 <KIR>
A:Cross-references: EMBL:U46839; NID:g1228142; PIDN:AAC52722.1; PID:g1228143
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
structurally and functionally unrelated lipophilic antitumor drugs.
C:Genetics:
A:Gene: mdr2
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr
F:1-637/653-1276/Region: duplication
F:409-603/Domain: ATP-binding cassette homology <ABC1>
F:426-433/Region: nucleotide-binding motif A (P-loop)
F:550-554/Region: nucleotide-binding motif B
F:1049-1245/Domain: ATP-binding cassette homology <ABC2>
F:1067-1074/Region: nucleotide-binding motif A (P-loop)
F:1192-1196/Region: nucleotide-binding motif B
F:88_94/Binding site: carboxylate (Asn) (covalent) #status predicted
F:432/Binding site: ATP (Lys) #status predicted
F:1072/Binding site: ATP (Lys) #status predicted

Query Match 35.5%; Score 2387; DB 1; Length 1276;
Best Local Similarity 42.1%; Pred. No. 2.5e-136;
Matches 534; Conservative 212; Mismatches 458; Indels 64; Gaps 13;

QY 97 KVN-LSFGLMRYATKMDLLIWIISTICAIIMASTFORIML-YQISYDEFD----- 146
Db 33 KNLGLGLTLFRISDWQDLFMFLGLIMAIHSGSLPLMKIVFGEMTDKFDVDTGNFSLP 92
QY 147 -----ELTKNLYEYVLIGIEFVYVSVGVFIYNGEHATOKIRRYLLE 190
Db 93 VNFSLMLNPGRLIEEMKRYAYISGLGGVLYAVIYQVSEPTLAAGQIKIRKRFH 152
QY 191 SLIRONIGFDKLGAGEVTTTADTNLIQDISSEKVGTLTALAFVTAFTIAYVYK 250
Db 153 ALIRQEMGFIDIKGTTELMTRLTDVSKISEIGDKVGFQIAINFAGFIYIGFRGMK 212
QY 251 LALIGSSTVALVLMGGSOFTIKSKSDSYAGGVAAEYVSIINNAFAIGTDKL 310
Db 213 LTLVIMAIISPIILGLSTAAVAAKILSTFSEKELAAVAKAGVAAEPARITVIAFGONKE 272
QY 311 AKOEVHLDAAEKWGTKNQIVMGFMIGAFGLMYSNYGFMNGSRFLVDGAVDGLIT 370
Db 273 LERYKHLENAKKIGIKKRAISANISKGLAFLLIYASYALAFYGSTLYLSKETITGNAMT 332
QY 371 VMAILIGSFSLGAVSPNQAFTNVAAMAKITFGTIDROSPLDPSYNECKTLDHFEHTE 430
Db 333 VFPSLILIGAFSVGOAACPIDAFANARGAAYVIFDIIDNNPKIDPSERGHKRDNIKGNLE 392

QY 431 LRNVKHIYSPRYVYVNEVSLSMAGKTTALVPGSGSKSTVGVVEREYMPVKGTVLL 490
Db 393 FSDVHFSPYRANIKIKLGLNLTAKVSGQVVALVNSGCGSKSTVQLLRDYDPEKISL 452
QY 491 DGHIDKDLNRMRLROQISLVSQEPVLEFRTIYKINRHLIGTYENESSEKVELLENAA 550
Db 453 DGQIDNRFNRYCRLEIIGVVSQEPVLESTTIAINIRG-----RGVYTMDE-----TEKAV 503
QY 551 KMANADFTALPEGETYVNGRGFLLSGGOKORIAIARAVSDPKTILLDEATSLDTR 610
Db 504 KEMAVDFIMKLPQKRDVLVGDGADLSGGOKRIARALVKNPKILLDEATSLDRE 563
QY 611 SEGVVQALERRAAEGRTTYIAHRLSTKTAHNIYVLVNGKIAEGSTHDELVDGGAAYRK 670
Db 564 SEAEVQAAALDKAREGRTTYIAHRLSTIRNADVIAFGEDGVIEQSHSELMKEGIYER 623
QY 671 LVQAORINEOKADALDEADAEELTNADIAKITASSASDDLQKPTTIDRTGTHSVSS 730
Db 624 LVNMQTAGSQILSEFEVELSDERKAAGDVAP-----NGMKARIRFNSTKSLKS 672
QY 731 ALISKRPPET-----TPKSLTILKFNVSFNRPEIYMLIGLVPSVLAGGQPTQAVYL 784
Db 673 PHQNRIDEETNELDANVPVPSFLKVLK-----LNKTEPFYVGVTCALANGALQPAFSLI 728
QY 785 YAKAISTLSIPESQYKLRHADFWSLMFEVGGIIQFTQSTNGAFAVCSERLIRRRAS 844
Db 729 LSEIMAIIFG-PGDDAVK-OOCKNMFSLVFLGLVLSFEFFLGGTFEKGAEILTRRLS 786
QY 845 TARTILRODIAFPDEENSTGALSTFSTFEMHLSGVSVTLTGITLMSTTGAAITTA 904
Db 787 MAFKAMLRQDMWFDHKSSTGALSTRLADAQVAGAGCTKALLAQNANAGTGTIIS 846
QY 905 LAIGMKLALVCSIVSVVLLACGFYRMYLAOFOSRKLAYEGSANFACENTSSIRTVASL 964
Db 847 FIYGMQDLTLLLSLVPEFIYAVAGIVEKMKLAGNAKRDKKEAAGKATATATEIRIVVSL 906
QY 965 TRRDVWEIYHAQDQGRSLISVRLSSILYASSQALVFCVALQFWYGTLL--GHHE 1022
Db 907 TORRKEFSMVEKTLHPYRNSVKRAHIYGTFSISQAFMYFSYAGCFRFGSYLIVNGHMK 966
QY 1023 Y-DIFFEFCFSEIILGAGSAGTVFSEFAPDMGAKNAAEFRRLPFRKQIDOWMSEGRK 1081
Db 967 FKVY---ILVFSALIVGAAVAGHASSFAFDYAKAKLSAALVSLFRKPLIDYSTSEGLM 1023
QY 1082 LETVEGEIEFRNRYFPTPRPEQVLRGLDLTVKPGQYVALVPGSGCGKSTYALIEREY 1141
Db 1024 PKDFEGSVTFNEVYFVPTPRANVPVLQGLSLVEKKQGTALVAGSSCGKSTVQLLIEREY 1083
QY 1142 DALAGSLVDGADISKLINSTRSEFLSVSQEPVLYQGITIKENILIGIYEDVDPEEFLIK 1201
Db 1084 DPMAGSVLLDDGDAKRLNQWMLAOLGIVSQEPVLYDSCIAENIAYAGDNSRVVPHDEIYR 1143
QY 1202 ACKDANIYDFINSLPGCFMNVVSGKGMLSGGOKOVATARALIRPKILLIDENASLD 1261
Db 1144 AKKANIHPIETLPQKRYTRVGDKTQLSGGOKRIARALIRPKILLIDENASLD 1203
QY 1262 SSESERVQALDAARGRITIAVHRLSTIRKADVIYVPEQGIYVSGHSELYOKRGYR 1321
Db 1204 TSESEKVFQALDKARSGRCIYIAHRLSTIQNADLVIVENKGVKRGHQQDLAQKGIY 1263
QY 1322 YELVNIQS 1329
Db 1264 FSNVNTQA 1271

RESULT 11

DVH03
multidrug resistance protein 3 - human
N:Alternate names: P-glycoprotein MDR3
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: J50051; S01346; A42213; T38015

DB 1271 SVQA 1274

RESULT 12

S41646

P-glycoprotein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001

C:Accession: S41646; S22354

R:Brown, P.C.; Thorngeltrson, S.S.; Silverman, J.A.

Nucleic Acids Res. 21, 3885-3891, 1993

A:Title: Cloning and regulation of the rat md2 gene.

A:Accession number: S41646; MUID:93376516; PMID:8103593

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1278

A:Cross-references: EMBL:L15079; NID:g310192; PIDN:AAA02937.1; PID:g310193

R:Deuchars, K.L.; Duthie, M.; Ling, V.

Biochim. Biophys. Acta 1130, 157-165, 1992

A:Title: Identification of distinct P-glycoprotein gene sequences in rat.

A:Reference number: S22351; MUID:92223089; PMID:1348630

A:Accession: S22354

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1211-1278 <DEU>

A:Cross-references: EMBL:X61105; NID:g56892; PIDN:CAA43417.1; PID:g1334220

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop

F:409-603/Domain: ATP-binding cassette homology <ABC1>

F:426-433/Region: nucleotide-binding motif A (P-loop)

F:1051-1447/Domain: ATP-binding cassette homology <ABC2>

F:1068-1075/Region: nucleotide-binding motif A (P-loop)

Query Match 35.1%; Score 2357; DB 2; Length 1278;

Best Local Similarity 40.9%; Pred. No. 1.6e-134;

Matches 524; Conservative 224; Mismatches 464; Indels 68; Gaps 14;

DB 89 LKTOLEIRYVNIISFFG---LMRYATKMDILIMVISTICAIAASTFORIML-YQISYDEF 144

89 LKTOLEIRYVNIISFFG---LMRYATKMDILIMVISTICAIAASTFORIML-YQISYDEF 144

DB 23 ISNQSSEKKKKVNLIGPLTLFRSDMODKLFMLGTMAMIAHSGSLPIMAVYGEKTDKF 82

23 ISNQSSEKKKKVNLIGPLTLFRSDMODKLFMLGTMAMIAHSGSLPIMAVYGEKTDKF 82

DB 145 YD-----ELTKNLYEVLIGIEFVTVYVSTVGFIYGEHA 180

145 YD-----ELTKNLYEVLIGIEFVTVYVSTVGFIYGEHA 180

DB 83 VNAGNFSPVNFPSLMLNGBRIIEBEEMTRVAYYSGLGCVLAIAYOVSEFTLAAAGRO 142

83 VNAGNFSPVNFPSLMLNGBRIIEBEEMTRVAYYSGLGCVLAIAYOVSEFTLAAAGRO 142

DB 181 TQKIRRYVLESTIRONIGFEDKIGAGEVTRITADNLTODGSEKVGLTITLAFVTA 240

181 TQKIRRYVLESTIRONIGFEDKIGAGEVTRITADNLTODGSEKVGLTITLAFVTA 240

DB 143 IRTIRKFFHAILROEMGFMDIKGTTELNTRLDDISIKISEGIGDVGMFFQAIATFFAG 202

143 IRTIRKFFHAILROEMGFMDIKGTTELNTRLDDISIKISEGIGDVGMFFQAIATFFAG 202

DB 241 FIAYVYKWKALICSTVALVLTWGGGSGQFIKYSKKSLSYAGAGTAAEVEVISIRN 300

241 FIAYVYKWKALICSTVALVLTWGGGSGQFIKYSKKSLSYAGAGTAAEVEVISIRN 300

DB 203 FIVGFIRGKRLTYIMATILGLSTAVAKILSTFSDKELAYAAAGAAVEALCAIRI 262

203 FIVGFIRGKRLTYIMATILGLSTAVAKILSTFSDKELAYAAAGAAVEALCAIRI 262

DB 301 AFAFGTQDKLAQYVHLDAERKWTGKNQIVMGFMIGAMFGIMYSNYGLGFMGSRFLVD 360

301 AFAFGTQDKLAQYVHLDAERKWTGKNQIVMGFMIGAMFGIMYSNYGLGFMGSRFLVD 360

DB 263 VIAFGQNKLEERYQKHLNNAKIKIGIKKAIISMGISIFLLIYASVALAFWVGSTLVIS 322

263 VIAFGQNKLEERYQKHLNNAKIKIGIKKAIISMGISIFLLIYASVALAFWVGSTLVIS 322

DB 361 GAVDVGDIITVLMALIGSFSLGAVSPNMAAFNAAAKITGTIDROSPLDPYNEGR 420

361 GAVDVGDIITVLMALIGSFSLGAVSPNMAAFNAAAKITGTIDROSPLDPYNEGR 420

DB 323 KEYTIGNATVFFSILIGAFVGOAPCIDAAPPNARGAAYVIFDIIDNPKIDSFSEERH 382

323 KEYTIGNATVFFSILIGAFVGOAPCIDAAPPNARGAAYVIFDIIDNPKIDSFSEERH 382

DB 421 TLDFHEGELRLNVKHIYSRPVTVMEVDSLSMPAGKTTALVPSGSGSTVYGLVER 480

421 TLDFHEGELRLNVKHIYSRPVTVMEVDSLSMPAGKTTALVPSGSGSTVYGLVER 480

DB 383 KPDSTIGNLEFSDVHFSYSRKANIKILKGLNLYKSGQVVALVNGSGCGKSTTVOLQRI 442

383 KPDSTIGNLEFSDVHFSYSRKANIKILKGLNLYKSGQVVALVNGSGCGKSTTVOLQRI 442

DB 481 YMFPGSTVLIDGHDIKDLNLRMLROQISLVSEPVLEGGTITTKIRHGLIGTKYENESD 540

481 YMFPGSTVLIDGHDIKDLNLRMLROQISLVSEPVLEGGTITTKIRHGLIGTKYENESD 540

DB 443 YDTEGTISIDGDDINRFNVRCLEFIIGVSGDEPVEFTTIANINRYG-----RGAVTMD 497

443 YDTEGTISIDGDDINRFNVRCLEFIIGVSGDEPVEFTTIANINRYG-----RGAVTMD 497

DB 541 KYRELLENAAKANAHDFTTALPEGETVVGORGFLSSGQKORIALAAVAVSDPKITLL 600

541 KYRELLENAAKANAHDFTTALPEGETVVGORGFLSSGQKORIALAAVAVSDPKITLL 600

DB 498 EIKK---AVKENAAYDFIMKLPQKEDTLVGDRGAQLSGGOKORIALAVALVNPKITLL 553

498 EIKK---AVKENAAYDFIMKLPQKEDTLVGDRGAQLSGGOKORIALAVALVNPKITLL 553

DB 601 DEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 660

601 DEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 660

DB 554 DEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 613

554 DEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 613

DB 661 LVDRGAVRKLVEAQRINQKADALEDAEDLTADNADIAKIKTASASSDDIGKPTTD 720

661 LVDRGAVRKLVEAQRINQKADALEDAEDLTADNADIAKIKTASASSDDIGKPTTD 720

DB 614 LIKKEGIVRLVLMQSSQIISEEPE-----VELSDEKAAGVAP--NGMKARIF 662

614 LIKKEGIVRLVLMQSSQIISEEPE-----VELSDEKAAGVAP--NGMKARIF 662

DB 721 RCTHNSVSSA-----ILSKRPETTPKYSKMLTLKFVASFNPPELPYMLIGVEY 772

721 RCTHNSVSSA-----ILSKRPETTPKYSKMLTLKFVASFNPPELPYMLIGVEY 772

DB 663 RNSTKSLKSSRAHQRLDVENELDANVPVSLKVLK---LNTKTEPFPVYGLCAI 718

663 RNSTKSLKSSRAHQRLDVENELDANVPVSLKVLK---LNTKTEPFPVYGLCAI 718

DB 773 LAGGQPTQAVLYAKAISTLSLPESQSKLRHADPWSLMFFVVGIIQIOTSTNGAANA 832

773 LAGGQPTQAVLYAKAISTLSLPESQSKLRHADPWSLMFFVVGIIQIOTSTNGAANA 832

DB 719 ANCALOPAFSIIISEMIALPG-PGDDTVK--QOKCNFSLVFLGIVHSFTTFLQCFITG 776

719 ANCALOPAFSIIISEMIALPG-PGDDTVK--QOKCNFSLVFLGIVHSFTTFLQCFITG 776

DB 833 VCSERLIRARSTAFRTILRODIAFPDKENSTGALTSELTETKLSGVSGVLTITLM 892

833 VCSERLIRARSTAFRTILRODIAFPDKENSTGALTSELTETKLSGVSGVLTITLM 892

DB 777 KAGEILTTLRSMAPKAMLRQDSWFDKHNSTGALSTRLANDPAQVQATGTRLLALINQ 836

777 KAGEILTTLRSMAPKAMLRQDSWFDKHNSTGALSTRLANDPAQVQATGTRLLALINQ 836

DB 893 TSTTLGAAIITIALATGKIALVCISVYPVLLACGFRTFMYLAQFQSKRLATEGSANPC 952

893 TSTTLGAAIITIALATGKIALVCISVYPVLLACGFRTFMYLAQFQSKRLATEGSANPC 952

DB 837 NTANIGTGIIISFVGMOTLTLILSVPRIVAVAGIYEMKMLGNANRDKREMAKIKAT 896

837 NTANIGTGIIISFVGMOTLTLILSVPRIVAVAGIYEMKMLGNANRDKREMAKIKAT 896

DB 953 EATSIIRTVASLTREDDWEIYHAQDQAGRTSLISVLRSSLLYASSQALVFCVAGFW 1012

953 EATSIIRTVASLTREDDWEIYHAQDQAGRTSLISVLRSSLLYASSQALVFCVAGFW 1012

DB 897 EALENIRTVASLTREDDWEIYHAQDQAGRTSLISVLRSSLLYASSQALVFCVAGFW 956

897 EALENIRTVASLTREDDWEIYHAQDQAGRTSLISVLRSSLLYASSQALVFCVAGFW 956

DB 1013 YGGLTL--GHHEV--DIREFVCESEILFQAQSGATVSEAPDMGKAKNAAEFRRLPDK 1069

1013 YGGLTL--GHHEV--DIREFVCESEILFQAQSGATVSEAPDMGKAKNAAEFRRLPDK 1069

DB 957 FGSYLLVNHMHMFQKV---ILVFSALIVLGAVALGHASAPADYAKAKLSAAVLFSLERQ 1013

957 FGSYLLVNHMHMFQKV---ILVFSALIVLGAVALGHASAPADYAKAKLSAAVLFSLERQ 1013

DB 1070 POIDNMESEGEKLEVEGEIEEPNNHARYPTREQVULNGDLTVRPGQVVALVYSGCGG 1129

1070 POIDNMESEGEKLEVEGEIEEPNNHARYPTREQVULNGDLTVRPGQVVALVYSGCGG 1129

DB 1014 PLTDYSIRSGMMPDKFEGSVTEFNEVFNPTRVANNVVLQGLSLVKKGGTLLVYSGSSGG 1073

1014 PLTDYSIRSGMMPDKFEGSVTEFNEVFNPTRVANNVVLQGLSLVKKGGTLLVYSGSSGG 1073

DB 1130 KSTTIALERFYDAIGSLIVDGKDISKLNINSYRSFLSLVSOEPLTYGTTIKENTLLGI 1189

1130 KSTTIALERFYDAIGSLIVDGKDISKLNINSYRSFLSLVSOEPLTYGTTIKENTLLGI 1189

DB 1074 KSTVVOGLERFYDPMAGYVLLDQGEAKKLVNQMLRQGLDISEPILTFPCSTAKNATYD 1133

1074 KSTVVOGLERFYDPMAGYVLLDQGEAKKLVNQMLRQGLDISEPILTFPCSTAKNATYD 1133

DB 1190 VEDDVEEFLIKAKDANIIDYFMSLPBEGNRYVSGKGMLSGGOKQVAVAIARALLRDK 1249

1190 VEDDVEEFLIKAKDANIIDYFMSLPBEGNRYVSGKGMLSGGOKQVAVAIARALLRDK 1249

DB 1134 NSRVVSGODEIVRAKKAHNIHPIETLPQKYEFTVSGDKTGLSGGOKQVAVAIARALLRDK 1193

1134 NSRVVSGODEIVRAKKAHNIHPIETLPQKYEFTVSGDKTGLSGGOKQVAVAIARALLRDK 1193

DB 1250 ILLDDEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 1309

1250 ILLDDEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 1309

DB 1194 VLLDDEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 1253

1194 VLLDDEATSAIDTRKSEGVQAAERAAEGRTTIVIAHRLSTIKTAHIVYLVNGKIAEGOTHE 1253

DB 1310 TSEELVQKKRYEYELVNLQS 1329

1310 TSEELVQKKRYEYELVNLQS 1329

DB 1254 THQQLAKQGIYFSWNIQA 1273

1254 THQQLAKQGIYFSWNIQA 1273

RESULT 13

148123

P-glycoprotein isoform III - Chinese hamster

C:Species: Cricetus griseus (Chinese hamster)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Feb-2001

C:Accession: 148123

R:Endicott, J.A.; Sarang, F.; Ling, V.

DNA Seq. 2, 89-101, 1991

A:Title: Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene fam

A:Reference number: 148121; MUID:92135896; PMID:1685679

A:Accession: 148123

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1281 <RES>

A:Cross-references: GB:M60042; NID:g191168; PIDN:AAA68885.1; PID:g191169

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop

F:412-606/Domain: ATP-binding cassette homology <ABC1>

Db	242	KSQGTVALVNGSGCKSTYQVLQMRLYLDPDLGWNVSIDGDIPTINVRILREILGVNGSR	301
Qy	515	VLESGTITKKNRHGLICTKYENESDEKVRRELIENAKMANAHDEFTALPEGETYNVGRG	574
Db	302	VLFPATTIAENIRYQ-----REDVTMDELEKAEKANAADFIMKLPHOFTDLYVBGRG	352
Qy	575	FLISGGKORIALARAVNSPKTILLIDDEATSAIDPSEBVOVALERAEGGTTIVIAHR	634
Db	353	AHVSGGKORIALARALVRNPKTILLIDDEATSAIDTSEBVAVALADKAEGETTIVIAHR	412
Qy	635	LSPTKTAHNIYVAVNGKILIAOGTHDELYBGCARFLVBAORINROKEADLEDAEDL	694
Db	413	LSYTRNNDVIAGFPGCGIYVQGNHDELMRKGIYFLVMTQVQAGNIELGNACKSKBEI	472
Qy	695	TNADIAKITASSASSDLDCKPTTIDRTGTHKSVSAILSKRPETTPYISLWTLKFEVA	754
Db	473	DNLMS---SKDSGSSSLIRRRSTRKSLCIGPHDDKRLSKREALDEBVPASFWRLTK---	526
Qy	755	SFNPRLPYMLIGIVFVSLVLAGGGOPVQVLYAKAISTLS---LPSPQSKLRHMDPMSL	811
Db	537	LSNSTEMPYFVAVGTECGLINGLQPPVPSYFVSXVGVFNGGPPETQ---RQNSNFTSL	581
Qy	812	MEFVAGIIOFTOSTNCAAAVAVCSERLIRARSTARFTILRDIAFFOKEENSTALTSF	871
Db	582	LEFLIGIISFTFPLQFTGCKAGELLITRLKRLRMVKSMLRODVSFMDPKNTTALTTR	641
Qy	872	LSTETKHLISGVSGVTLGTILMTSTYGAALIIALAIQMKLALVCISVAVPVLACGEYFY	931
Db	642	LANOAAOVKCATGATGRILAVIIONIANLGTGIIISLIGMOLTLLALIVAIYIAIAGVEMK	701
Qy	932	MLAOFQSKSLAYGSSANFCAEPAATSSIRYASLITRDRDWEIYHQLDAOGTSLISVLR	991
Db	702	MLSGOALDKKEELGSSKIAITEALIEENRIVYSLTRQKEETMYASLOLIPYNNAKKAV	761
Qy	992	SSLIVASSOALVPEFCVALLGFWYGTLLLGHHNEYDIFRPFCESEILFGAOSACTVESPAD	1051
Db	762	FGIIFSTQAMMYRYAACAIRFPAJLVYQOLMFHEVNLVFSALYFGMAAAOVQVSPAD	821
Qy	1052	MGRKKNAAEERFLRDLRKPQIDIMWSEBGEKLETVBGEIEFRNVHRRPTREPOVYLGLD	1111
Db	822	YAKATVASASHIRIETKPEPIDYSTOGLKPNMLBENVOFSGVVPYNPYPRPSIPVLOGLS	881
Qy	1112	LTVAPQVVALVNGSGGCKSTYQVLQMRLYLDPDLGWNVSIDGDIPTINVRILREILGV	1171
Db	882	LEVKGQGLTALVNGSGGCKSTYQVLQMRLYLDPDLGWNVSIDGDIPTINVRILREILGV	941
Qy	1172	QEPFLYOGTITKENILIGIVEDVPEEFLKACDAKIATYDFINSLDEGEFTVNGSGKMLS	1231
Db	942	QEPFLPFCSTLAENYAYDNRVSYEIEIYRAAKENIHPIDTSLDKYNTNRGDKGLOLS	1001
Qy	1232	GGKORVAILARALIRDBKILLDDEATSAIDSESEKVVQALDAARGRTTIAVAHRSTI	1291
Db	1002	GGKORVAILARALVROPHILLIDDEATSAIDTSESEKVVQALDKAREGRTCIIVAHIRSTI	1061
Qy	1292	OKADIVIVPOGKRVESGTSESLYQKGRYIELVNLQSLGK 1332	
Db	1062	ONADLIVIVIONGKVEKHEGTHQOLLACKGIFYSVSYOAAK 1102	

RESULT 15
 T422228
 P-glycoprotein sister - rat
 N:Alternate names: multidrug resistance protein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-Jul-2002
 C:Accession: T422228
 R:Childs, S.J.; Yen, R.L.; Hul, D.; Ling, V.
 submitted to the EMBL Data Library, June 1997
 A:Description: Taxol resistance mediated by the liver-specific Sister gene of P-glycoprotein
 A:Reference number: Z22088
 A:Accession: T422228
 A:Status: preliminary; translated from GB/EMBL/DBJ

[illegible]

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:51:33 ; Search time 30 Seconds

(without alignments)
1844.314 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLTNPISLPETAMREPAE.....YOKGRYELVNLQSLGKH 1334

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3068.5	45.7	1362	1 PMD1_SCHPO	P36619 schizosacch
2	2493	37.1	1276	1 MDR1_CRIGR	P21448 cricetulus
3	2489	37.0	1280	1 MDR1_HUMAN	P08183 homo sapien
4	2471.5	36.8	1276	1 MDR3_MOUSE	P21447 mus musculu
5	2453.5	36.5	1276	1 MDR1_MOUSE	P06795 mus musculu
6	2413	35.9	1276	1 MDR2_CRIGR	P21449 cricetulus
7	2392.5	35.6	1277	1 MDR1_RAT	P43245 rattus norv
8	2367	35.5	1276	1 MDR2_MOUSE	P21440 mus musculu
9	2379.5	35.4	1279	1 MDR3_HUMAN	P21439 homo sapien
10	2357	35.1	1278	1 MDR2_RAT	P08201 rattus norv
11	2357	35.1	1281	1 MDR3_CRIGR	P23174 cricetulus
12	2328	34.6	1321	1 AB11_HUMAN	O95342 homo sapien
13	2247	33.4	1321	1 AB11_MOUSE	O93430 mus musculu
14	2235	33.3	1321	1 AB11_RABIT	O93430 mus musculu
15	2230.5	33.2	1321	1 AB11_RAT	O93430 mus musculu
16	1985.5	29.5	1321	1 MDR1_MOUSE	P21447 mus musculu
17	1971	29.3	1302	1 MDR4_MOUSE	P34712 caenorhabd
18	1886.5	28.1	1254	1 MDR3_MOUSE	P00449 drosophila
19	1877.5	27.9	1302	1 MDR5_MOUSE	P34713 caenorhabd
20	1745	26.0	1280	1 MDR1_MOUSE	O00748 drosophila
21	1432.5	21.3	1419	1 MDR1_MOUSE	O06034 leishmania
22	1229.5	18.3	1290	1 MDR1_MOUSE	P13568 plasmodium
23	956	14.2	1336	1 MDR1_MOUSE	P12866 saccharomyc
24	927	13.8	715	1 AB10_MOUSE	P78966 schizosacch
25	890	13.5	738	1 AB10_MOUSE	O93139 mus musculu
26	842	12.5	735	1 AB10_MOUSE	O93139 mus musculu
27	830	12.3	1323	1 HST6_CANAL	O93139 mus musculu
28	820.5	12.2	762	1 ABC9_MOUSE	O93139 mus musculu
29	817.5	12.2	762	1 ABC9_MOUSE	O93139 mus musculu
30	814.5	12.1	766	1 ABC9_MOUSE	O93139 mus musculu
31	749.5	11.2	582	1 MSBA_ECOLI	P27239 escherichia
32	746.5	11.1	1325	1 MRP4_HUMAN	O15439 homo sapien
33	743	11.1	1502	1 MRP4_HUMAN	O88269 rattus norv

34	727	10.8	685	1 MDL1_CANAL	P97998 candida alb
35	724	10.8	1522	1 MRP3_RAT	O88563 rattus norv
36	707.5	10.5	1743	1 TAGC_DICDI	O23868 dictyosteli
37	706.5	10.5	1503	1 MRP6_HUMAN	O95255 homo sapien
38	698	10.4	725	1 TAP1_RAT	P36370 rattus norv
39	693	10.3	607	1 HEPA_ANASP	P22638 anabena sp
40	692	10.3	1527	1 MRP3_HUMAN	O15438 homo sapien
41	689.5	10.3	724	1 TAP1_MOUSE	P21958 mus musculu
42	687	10.2	584	1 LMRA_LACTA	O93108 lactococcus
43	684.5	10.2	1515	1 YCFL_YEAST	P39109 saccharomyc
44	679	10.1	1436	1 MRP5_MOUSE	O9715 mus musculu
45	678	10.1	1592	1 YHDS_YEAST	P38735 saccharomyc

ALIGNMENTS

RESULT 1
PMD1_SCHPO STANDARD; PRT; 1362 AA.
AC P36619; 074513;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leptomycin B resistance protein pm1.
GN PMD1 OR SPCC663.03.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX MEDLINE=92244046; PubMed=1349418;
RA Nishi K., Yoshida M., Nishimura M., Nishiyama M.,
RA Hironouchi S., Bepu T.;
RT "A leptomycin B resistance gene of Schizosaccharomycetes pombe encodes
RT a protein similar to the mammalian P-glycoproteins.";
RL Mol. Microbiol. 6:761-769(1992).
RN [2]
RP SEQUENCE FROM N.A.
RP SPRAIN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Gellies S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
RA Skellton J., Simmonds M., Squares R., Squares S., Sharp S.,
RA Taylor K., Taylor R.G., Riley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickberg G., Aert R., Roben J., Gymnopoulos B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzjak K., Langer I., Beck A., Leirich H., Reinhardt J., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lalauze V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas R., Rochez M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovsky G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:811-880(2002).
CC -I- FUNCTION: MAY BE A TRANSMEMBRANE TRANSPORTER OF THE MATING
CC FACTOR, NAMELY P-FACTOR OR M-FACTOR. CONFERS RESISTANCE TO
CC LEPTOMYCIN B AND TO SEVERAL OTHER ANTIFUNGAL DRUGS.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

```

-! SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).

-----
CC CC      EMBL; D10695; BAA01537.1; -.
CC CC      DR EMBL; AL031307; CAA20363.1; -.
CC CC      DR PIR; S20548; S20548.
CC CC      DR InterPro; IPR003593; AAA_Atpase.
CC CC      DR InterPro; IPR003439; ABC_transportr.
CC CC      DR InterPro; IPR001140; ABCTransplfTW.
CC CC      DR Pfam; PF00005; ABC_tran; 2.
CC CC      DR Pfam; PF00664; ABC_membrane; 2.
CC CC      DR ProDom; PD000006; ABC_transporter; 2.
CC CC      DR SMART; SM00382; AAA; 2.
CC CC      DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
CC CC      KW Transmembrane; ATP-binding; Glycoprotein; Antibiotic resistance.
KW DOMAIN
FT TRANSMEM          92   115       POTENTIAL.
FT TRANSMEM          138   162       POTENTIAL.
FT TRANSMEM          220   237       POTENTIAL.
FT TRANSMEM          244   264       POTENTIAL.
FT TRANSMEM          320   346       POTENTIAL.
FT TRANSMEM          354   374       POTENTIAL.
FT DOMAIN            375   788       CYTOPLASMIC (POTENTIAL).
FT TRANSMEM          789   809       POTENTIAL.
FT TRANSMEM          835   859       POTENTIAL.
FT TRANSMEM          916   935       POTENTIAL.
FT TRANSMEM          940   957       POTENTIAL.
FT TRANSMEM          1022  1040       POTENTIAL.
FT TRANSMEM          1054  1072       POTENTIAL.
FT DOMAIN            1073  1362       CYTOPLASMIC (POTENTIAL).
FT NP_BIND           1073  1362       ATP (POTENTIAL).
FT NP_BIND           1154  1161       ATP (POTENTIAL).
FT CONFLICT          1334  1335       TA -> TC (IN REF. 1).
SQ SEQUENCE          1362 AA; 149631 MW; 2ADFBABE7B3ACEEB CRC64;

Query Match              45.7%; Score 3068.5; DB 1; Length 1362;
Best Local Similarity    47.3%; Pred. No. 2.9e-173;
Matches 656; Conservative 233; Mismatches 403; Indels 95; Gaps 18;

OY 14 AMREPAETSTTEEOASTPHADEKRLSLDLSAPSTATPADKEHPKSSSNNAVSYNE 73
   :: :|:: |:: ||:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
DB 2 SLHKSRSYSTVMD---NHSLD----LKIKLPSEKKNFSTEKR--NEASESHVDYVK 52

OY 74 VDALIAHLPEDEROYLKTQLEEKIVNISFF GLMRYATKMIDLINVISTICAIAA---- 127
   : : :|:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
DB 53 -DPFEQVPYE-EQELIYLKOINDPTAKLTGYPRIISLYADKWDIMQLAGITIGIGALGP 110

OY 128 -----ASTQRRLMLQOISIDERYDELTKRNVLYFYVLGLIGEFPYTV 166
   ||::||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||::
DB 111 LMSLVSGOLAQAFDTDLASGSKASSFOH-----TYDHF-----CLFYIIAIGVEGCS 157

OY 167 YVSVTGFYTGEHAHQTRREYLEESLRONIGYEPDKLAGEVTRYTRTDNLIODGISRK 226
   |:|:||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||::
DB 158 YYYTTFIAGRIARRRRÖDYLAHLAISONTNGYPRLGAGEITRTITTDTNFIDDGLEK 217

OY 227 VGLTLTALATEVTAFLIAYVRWKLLICSSTVALVLFMGGSQSOFITIKYSKSLDSYG 286
   ||::||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||::
DB 218 VGLVEFAATLEVFSGVFIARIHKMKFTLLISSMFPAIGCGTGIGVDFITKNKGAIYVAE 277

OY 287 GGTVAEVIYSIRRNATACGTODKLAKÖYEHVLIDEKKKGTTNQÖYMGMICAMGLMTSN 346
   ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||::
DB 278 SSTFEEEVFNSNRNAFGTÖDIILAKTNKYLIITAORGIKKAIALMGVMFMFEVAVGY 337

OY 347 YGLGFWMASRFYDOCAVVVDILVYTMALMLIGSSLVGSSLNVAOAFTNAVANAAARKTGTI 406
   ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||::
DB 338 YGLAFWEGGRILLHADDDLVSKLIGCFPAVLLASVLAISIPMOSESFCSCAAARKITEPTI 397
```

[illegible]

```

AC P21448;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein 1).
OS ABCB1 OR PGY1 OR PGP1.
GN Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetidae.
XX NCBI_TaxID=10029;
RN NCB1.
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangl F., Ling V.;
RT "complete cDNA sequences encoding the Chinese hamster P-glycoprotein
   gene family."
RL DNA Seq. 2:89-101(1991).
RN NCB1.
RP SEQUENCE FROM N.A.
RX MEDLINE=91154265; PubMed=1671863;
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced pgp1 transcripts in multidrug-
   resistant Chinese hamster lung cells."
RL J. Biol. Chem. 266:4545-4555(1991).
RN NCB1.
RP SEQUENCE OF 706-1276 FROM N.A.
RX MEDLINE=8812213; PubMed=2893255;
RA Endicott J.A., Turanka P.F., Sarangl F., Gerlach J.H., Deuchars K.L.,
   Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
   drug-sensitive Chinese hamster ovary cells."
RL Mol. Cell. Biol. 7:4075-4081(1987).
CC -1. FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
   DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1. SUBCELLULAR LOCATION: Integral membrane protein.
CC -1. MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
   CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
   CANNOT.
CC -1. SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M60040; AAA68883.1; -.
CC DR EMBL; M59253; AAA37004.1; -.
CC DR EMBL; M17897; AAA37006.1; -.
CC DR PIR; A38696; DVHYIC.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR003439; ABC_transportr.
CC DR InterPro; IPR001140; ABCtransplrtm.
CC DR Pfam; PF00005; ABC_tran; 2.
CC DR Pfam; PF00664; ABC_membrane; 2.
CC DR ProDom; PD00006; ABC_transportr; 2.
CC DR SMART; SM00382; AAA; 2.
CC DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
CC KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
   Multigene family.
CC FT DOMAIN 1 50 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 51 71 POTENTIAL.
CC FT TRANSMEM 117 137 POTENTIAL.
CC FT TRANSMEM 186 206 POTENTIAL.
CC FT TRANSMEM 213 233 POTENTIAL.
CC FT TRANSMEM 294 314 POTENTIAL.
CC FT TRANSMEM 323 343 POTENTIAL.
CC FT DOMAIN 344 707 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 708 728 POTENTIAL.
CC FT TRANSMEM 754 774 POTENTIAL.

```

[illegible]

DB 846 ISLIIGWLTLLIATVPIIATAGVEMKLSQALDKKKEGSKITATELENTRYV 905
QY 963 SLTREDDVMEIYHAQDAQRTSLISVLRSLLYASSQALVEFCVAGFWYGGTLGHHB 1022
DB 906 SLTRBQKFNMYAQSIIQIYRNALKAHVGITFSFTQAMMFYSVAACFRGAVLYAREL 965
QY 1023 YQIFRFVCFSELPGACAGVYFSAFPMGKAKNAAEFRRLFDKRPQIDMKSEGEYL 1082
DB 966 MTFENVLVFSAIVFGAMVGVSSFPADYAKAKVSASHIIMEIEVPSIDSTGTGLRP 1025
QY 1083 EYVEGEIERFNRHFRYPRPEQPVLRGLDLYVKGQYVHLYVPGSGCKSTTALREFYD 1142
DB 1026 NLEBGVNRKNEVYENFTPRDIPVIOGLMEYKKGOTIALVSSGCKSTVVOLEFRFD 1085
QY 1143 AIAIGSLVDKDISKILINSYFSLSVQOEPTLYOGTICKENILIGIVEDDVEEFLIRA 1202
DB 1086 PMAGFVFLDGKEVQNLQVOMLRHLGIVSQEPILFDCSIENIAYDNGSRVYSQDEIERA 1145
QY 1203 CDANITVDYFINSLEPFNFVYSGKGLSGGOKORVATRALRLDKIILLDEATSAIDS 1262
DB 1146 AAEANIHQIESLPDKYNTRVGDKGTQSLGOKORVATRALVROPHIILLDEATSAIDT 1205
QY 1263 ESEKVVQALDAARGTITIAVAHRLSTQKADVIYVFOGKIVESGTHSELVOKKGRY 1322
DB 1206 ESEKVVQALDAARGTITIAVAHRLSTQKADVIYVFOGKIVESGTHSELVOKKGRY 1265
QY 1323 ELVNTSLQK 1332
DB 1266 SMVSQVQAK 1275

RESULT 3
MDRL_HUMAN
ID MDRL_HUMAN STANDARD; PRI: 1280 AA.
AC P08183; Q12755; Q14812;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein 1) (CD243 antigen).
GN ABCB1 OR PGT1 OR MDRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87028230; PubMed=2876781;
RA Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,
RA Robinson I.B.;
RT "Internal duplication and homology with bacterial transport proteins
in the mdrl (P-glycoprotein) gene from multidrug-resistant human
cells";
RL Cell 47:381-389(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90094448; PubMed=1967175;
RA Chen C.-J., Clark D.P., Ueda K., Pastan I., Gottesman M.M.,
RA Robinson I.B.;
RT "Genomic organization of the human multidrug resistance (MDRL) gene
and origin of P-glycoproteins";
RL J. Biol. Chem. 265:506-514(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97190336; PubMed=9038218;
RA Chen G., Duran G.E., Steger K.A., Lacayo N.J., Jaffrezou J.P.,
RA Dumontet C., Sikic B.I.;
RT "Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein,
altered phenotype, and resistance to cyclosporins";
RL J. Biol. Chem. 272:5974-5982(1997).
RN [4]
RP SEQUENCE OF 1-234 FROM N.A.
RA Smith A., Beck C., Gibson A.;

RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE OF 178-215 AND 800-856 FROM N.A.
RX MEDLINE=90290529; PubMed=1972623;
RA Gekeler V., Weger S., Probst H.;
RT "mdr1/P-glycoprotein gene segments analyzed from various human
leukemic cell lines exhibiting different multidrug resistance
profiles";
RL Biochem. Res. Commun. 169:796-802(1990).
RN [6]
RP SEQUENCE OF 1-23 FROM N.A.
RA Kioka N., Tsubota J., Kakehi Y., Komano T., Gottesman M.M.,
RA Pastan I., Ueda K.;
RL Submitted (JUL-1991) to the EMBL/Genbank/DBJ databases.
RN [7]
RP VARIANTS SER-893 AND THR-893.
RX MEDLINE=21686803; PubMed=11829140;
RA Saito S., Iida A., Sekine A., Mura Y., Ogawa C., Kawuchi S.,
RA Higuchi S., Nakamura Y.;
RT "Three hundred twenty-six genetic variations in genes encoding nine
members of ATP-binding cassette, subfamily B (ABCB/MDR/TRP), in the
Japanese population";
RL J. Hum. Genet. 47:38-50(2002).
CC - FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SMALL INTESTINE
CC AND BRAIN.
CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC - DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/PGT1ID105.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, M14758; AAA59575.1; -
DR EMBL, M29447; AAA59576.1; -
DR EMBL, M29447; AAA59576.1; JOINED.
DR EMBL, M29445; AAA59576.1; JOINED.
DR EMBL, M29426; AAA59576.1; JOINED.
DR EMBL, M29427; AAA59576.1; JOINED.
DR EMBL, M29428; AAA59576.1; JOINED.
DR EMBL, M29429; AAA59576.1; JOINED.
DR EMBL, M29430; AAA59576.1; JOINED.
DR EMBL, M29431; AAA59576.1; JOINED.
DR EMBL, M29432; AAA59576.1; JOINED.
DR EMBL, M29433; AAA59576.1; JOINED.
DR EMBL, M29434; AAA59576.1; JOINED.
DR EMBL, M29435; AAA59576.1; JOINED.
DR EMBL, M29436; AAA59576.1; JOINED.
DR EMBL, M29437; AAA59576.1; JOINED.
DR EMBL, M29438; AAA59576.1; JOINED.
DR EMBL, M29439; AAA59576.1; JOINED.
DR EMBL, M29440; AAA59576.1; JOINED.
DR EMBL, M29441; AAA59576.1; JOINED.
DR EMBL, M29442; AAA59576.1; JOINED.
DR EMBL, M29443; AAA59576.1; JOINED.
DR EMBL, M29444; AAA59576.1; JOINED.
DR EMBL, M29445; AAA59576.1; JOINED.
DR EMBL, M29446; AAA59576.1; JOINED.
DR EMBL, AF002457; AAB69423.1; -
DR EMBL, AF002457; AAB69423.1; -
DR EMBL, M37724; AAA88047.1; -
DR EMBL, M37725; AAA88048.1; -
DR EMBL, X58723; CAA41558.1; -
DR PIR: A25059; DVH01.
DR PIR: A34914; A34914.
DR Genew; HGNC:40; ABCB1.

DR MIM: 171050; AAA_ATPase.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABCtransport.
 DR Pfam: PF00005; ABC_tran. 2.
 DR Pfam: PF00664; ABC_membrane. 2.
 DR ProDom: PD000006; ABC_transport. 2.
 DR SMART: SM00382; AAA. 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family; Polymorphism.
 FT DOMAIN 1 51
 FT TRANSMEM 52 72 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT TRANSMEM 216 236 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 326 346 POTENTIAL.
 FT DOMAIN 347 710 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 711 731 POTENTIAL.
 FT TRANSMEM 757 777 POTENTIAL.
 FT TRANSMEM 833 853 POTENTIAL.
 FT TRANSMEM 854 874 POTENTIAL.
 FT TRANSMEM 937 957 POTENTIAL.
 FT TRANSMEM 974 994 POTENTIAL.
 FT DOMAIN 995 1280 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 427 434 ATP (BY SIMILARITY).
 FT NP_BIND 1070 1077 ATP (BY SIMILARITY).
 FT REPEAT 638 1280
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 893 893 A -> S (IN DBSNP:2032582).
 FT VARIANT 893 893 A -> T.
 FT CONFLICT 23 23 /FTID=VAR_013362.
 FT CONFLICT 185 185 S -> R (IN REF. 6).
 FT CONFLICT 336 336 G -> V (IN REF. 1 AND 3).
 FT CONFLICT 412 412 MISSING (IN REF. 3).
 FT CONFLICT 438 438 G -> A (IN REF. 3).
 FT CONFLICT 438 438 Q -> S (IN REF. 3).
 SQ SEQUENCE 1280 AA: 141462 MW: 4816795143675 CRC64;
 Query Match 37.0%; Score 2489; DB 1; Length 1280;
 Best Local Similarity 42.1%; Pred. No. 4, 5e-139;
 Matches 555; Conservative 229; Mismatches 448; Indels 60; Gaps 9;
 QY 94 EEIKVNISFGLMRVATKMDILIMVISTICAIAAASTFORIMLY----- 137
 DB 28 KKKKPTVVSFMRYSNMDLKLVMVVTLLAIIHGAGLPLMLVFGEMTDIFANAGMLED 87
 QY 138 -----QISDYDEY----DELTKNVLYEVLGICEFTVVSIVTGFITYGHEAOKI 184
 DB 88 LMSNITNSDINTGFFMNLSEEDMTTRAYYSIGAGVLAAYIYVSFWCLANGROTHKI 147
 QY 185 REYLESILRONIGYFDKLCAGEVTTITADTNLIODGISEKGLTFTALFVATFIIA 244
 DB 148 RKDFPFAIMQELIGWFDVHVGELNRLTLDVSKINIGIDKIGMFQSNATPTGTIGV 207
 QY 245 YKRYKALICSTIVALTVMGGSOFTIKYKKSIDSYGAGTVAAEVISSIRNATAF 304
 DB 208 FTRGMKLTITVILAIISPLGLSAVWMAKLISSFDKELLAVAKAGVAEEVLAIRVIATF 267
 QY 305 GPQDKLAKQVEVLDLAEKKGTKNQIYMGPMICAMGCLMSNGGLWMSRPLVDGAVD 364
 DB 268 GGGKKELERNNKLEKKRIGIKAITANISIGAAFLIYASALAWYGTITVLSGEYS 327
 QY 365 VGDILVYMAILIGSFLGNVSPNAQFTNAVAAAARIFGTIPROSLDIPYNSNGKTLHD 424
 DB 328 IGVLVVFEFVLLGAFSGVSGASPSIEFANAARAAIETIFIIIDKPSIDISYSGHAPDN 387
 QY 425 FEGHIELRNKHIYPSRPEVTYMEDVSLSMAPAKTALVGPSSGKSTVVGLEVFYMPV 484

DB 388 IKGNLEFRNVHFSYPSRKEKILKGLNLIKVOGGQVVALVNSGGCKSTVOQLMRDLP 447
 QY 485 RGVLLDGHIDKIDNLRLWLOQISLVSOEPLFETTYTKNRHLLICTKXENESEDVRE 544
 DB 448 EGVSVGDQDRTNVTNLEIGVSOEPLVLTAEINTRYG-----RENVMD----- 499
 QY 545 LIENAAKMANAHDTALPEGETYTNVSGRGLSGGOKRIARAAVSPKILLDEAT 604
 DB 500 -IEAAVEANAYDPIKMLPKHFDLVEBERGQALSGGOKRIARALVNPRIKILLDEAT 558
 QY 605 SALDTKEGVVQALERPAEGRTTIVAHRLSTIKTANHVLYNGKIAEGTHDELVD 664
 DB 559 SALDTEEAAYVQALDKARGRRTTIVAHRLSTVRNADVLAGFDGIVYKGNHDELMKE 618
 QY 665 GGAARKIVEORINEOKADALDEADDLTNMDIAIKTASSASDLDKPTIDTGT 724
 DB 619 KGIFKLVMTQAGNEVE--LENA-----ADESSEIDALEMSNDSSILRRST 668
 QY 725 HKVSAAILSKRP-----PETPKYSIMTLTKFVASFNRPEIPYMLIGLVSVLAGGG 777
 DB 669 RRSVRSQADRKISTKEALDESTIPVSPFRIMK-----INTEPPTVVGVCALINGCL 724
 QY 778 QPTQAVLYAARISTLSPEQYSKLRHDADEFWSLMFEVVGIIQPTQSTNGAFAVCSER 837
 DB 725 QPAPAIIFSRIKIGVTRIDDEPK-RQNSMLSELLFALIGIISPTIFELQGTGKAGEI 783
 QY 838 LIRARSTARTIIRQIAFFDKENSTGALTSLSLSTETHLSGVSGVTTGILMTSTL 897
 DB 784 LTKRLRVFRSMRLROVSNFDPKNTGALTTLRLADAQVKCAISRLAVITQNIANTL 843
 QY 898 GAATIIIALGKMLATCISVVPVLLACGFYRFYMLAOFQSRSLAEAGSANFACEATSS 957
 DB 844 CTGIIISFTYQWLTLLLAIVPIIALAGVEMKLSGALKKKELEGAKIATEAEN 903
 QY 958 IRTVASLFRERDWEIYHAOLDAGRTSLISVRSLLYASSQALVEFCVALGFWYGT 1017
 DB 904 FRFVVSITQDKFEMHMAQSLQVYRNSLRKAHIFGTFSTQMMTFVSAAGCFRCAYL 963
 QY 1018 LGHHEVYIFRFVFEFSLILGAGASGTVSPADPMGAKANAAEFRLPRDKPIDMWSE 1077
 DB 964 VAHKLMSFEDVLVFSVVGMAVAGVSPADPAKAKISNAITIMIETPTLIDYST 1023
 QY 1078 EGEKLTVEGEIEFRNVHFRTPRPOVLRGLDLYTKPGQVVALVGPSSGCKSTTALL 1137
 DB 1024 EGLMPTNLEGNVTGEVFNTPRPDIPLVLOGLSLEPKKGTLLVSSGCGKSTVQL 1083
 QY 1138 ERFYDAIAGSLVDGKDISKLNINSYRSLVSQSEPTLYOGITKENILLGIYEDVPEE 1197
 DB 1084 ERFYDPLAGKVLDDGKRIKRLNVQMLRAHLCIYVSEPTLFCDSIAENIAYGDNRRVSSOE 1143
 QY 1198 FLTKACDANITDPIFMSLPEGFNTVVGSKGMSGCKORAVATALLRDKILLDEAT 1257
 DB 1144 EIVRAAKENIHAIETSLPKYSTKVGDKTQLSGGOKRIARALVNPRIKILLDEAT 1203
 QY 1258 SALDSESEKYVQALDAAGRTTIVAHRLSTOKRADVIVYFPOGKIVSGHSELVOK 1317
 DB 1204 SALDTESEKYVQALDAAGRTTIVAHRLSTOKRADVIVYFPOGKIVSGHSELVOK 1263
 QY 1318 KGRYELVNLQS 1329
 DB 1264 KGIFYSMVSVOA 1275
 RESULT 4
 MDR3_MOUSE
 ID MDR3_MOUSE STANDARD; PRT; 1276 AA.
 AC P21447;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Multidrug resistance protein 3 (P-glycoprotein 3) (MDR1A).
 GN ABCB1A OR ABCB4 OR PGY3 OR PGY-3 OR MDR3 OR MDR1A.

Db 783 RLRYVFKSMKRODVSWEDDPKNTGALTTLRLANDAAQVKGAGSLAVIFONTANTG 842
 QY 901 IITAIAGMKALVCISVVPVLLACGFYFVYMLAQFQSRKLAEGSANACEATSSIRT 960
 Db 843 IITSLIYGMQLTLLALVYIITAGVYEMKMLSGQALKKKEEGSKATATEIEMFRT 902
 QY 961 VASLTREBDVWEIYHQAQDAGRTSLISVLRSSLLVASSQALVFCVYALGFYWGTLGH 1020
 Db 903 VVSLTRBQKFEYMAQSLQIPYRNAMKKAHVFGITFFPTQAMFYAAACFRGAYLVTO 962
 QY 1021 HEVDIRRFVCFSEILFGAASACTVPSFARDMGAKNAAEFRRLFRKQIDWMSSEGE 1080
 Db 963 QLMTFENVLLVFAIVGAAVQVSSFPDYAAATVSAASHIIRLIEKTEPIDISYTOGL 1022
 QY 1081 KLETFVEGEIEFRVHFRPTPEQPVYRGDLTVKPGQYVALVPSGCGSTTALLERF 1140
 Db 1023 KPMMLBGNVQSGFVFYFPRPSPVQGLSTLEKKKGOTLALVSSCGSTVYQLLERF 1082
 QY 1141 YDAIAGSILVDGDISKLNINSYRSFLSLVSOEPTLYOGTKEINILLGIYEDDVEEFLI 1200
 Db 1083 YDPMAGSVFLDGKEIKQLNQLVQLRAQGIYVQEPILFDCSIAENIAYGDSRVYSYEIV 1142
 QY 1201 KAKDANIVYFIMSLPGEFTVYVSGKGMLSGCKORVATARALLRDKILLDEATASL 1260
 Db 1143 RAKEANIHQFIDSLPRKYTPVGDGKTQSLGCKORHIALARALVROPHILLDEATASL 1202
 QY 1261 DSESEKVVQALDPAARFTTIAVHRLSTYOKADVIYVFDGKIVSGTSHSELVOKGR 1320
 Db 1203 DTSEKVVQALDKARGRCIYIAHRLSTIQNADLIVLQNGKVRKHGHQQLAOKGI 1262
 QY 1321 YVELVNTQSLGR 1332
 Db 1263 YFSMVSVOAGAK 1274

RESULT 5
 MDRL_MOUSE
 ID MDRL_MOUSE STANDARD; PRT; 1276 AA.
 AC P06795;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Multidrug resistance protein 1 (P-glycoprotein 1).
 GN ABCB1 OR ABCB1B OR PGY1 OR PGY1-1 OR MDRL1 OR MDRLB.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=87028229; PubMed=3768958;
 RA Gros P., Croop J., Housman D.;
 RT "Mammalian multidrug resistance gene: complete cDNA sequence
 indicates strong homology to bacterial transport proteins.";
 RL Cell 47:371-380(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=89367274; PubMed=2570420;
 RA Raymond M., Gros P.;
 RT "Mammalian multidrug-resistance gene: correlation of exon
 organization with structural domains and duplication of an ancestral
 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).
 RN [3]
 RP SEQUENCE OF 1-21 FROM N.A.
 RP MEDLINE=91042535; PubMed=2248681;
 RA Raymond M., Gros P.;
 RT "Cell-specific activity of cis-acting regulatory elements in the
 RT promoter of the mouse multidrug resistance gene mdrl.";
 RL Mol. Cell. Biol. 10:6036-6040(1990).
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
 CC LINKER DOMAIN.
 CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
 CC RELATED BUT DISTINCT CELLULAR GENES.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14757; AAA79005.1; -;
 DR EMBL: M60348; AAA39513.1; -;
 DR PIR: A33719; DVMSL
 DR MGI: 97568; Abcd1b.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABCtransport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family; Phosphorylation.
 FT DOMAIN 1 47
 FT TRANSMEM 48 71 POTENTIAL.
 FT TRANSMEM 119 139 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 296 316 POTENTIAL.
 FT TRANSMEM 327 347 POTENTIAL.
 FT DOMAIN 348 709 POTENTIAL.
 FT TRANSMEM 710 730 POTENTIAL.
 FT TRANSMEM 755 775 POTENTIAL.
 FT TRANSMEM 831 851 POTENTIAL.
 FT TRANSMEM 855 882 POTENTIAL.
 FT TRANSMEM 935 955 POTENTIAL.
 FT TRANSMEM 966 986 POTENTIAL.
 FT DOMAIN 987 1276 POTENTIAL.
 FT NP_BIND 426 433 ATP (BY SIMILARITY).
 FT NP_BIND 1068 1075 ATP (BY SIMILARITY).
 FT REPEAT 1 635
 FT CARBOHYD 636 1276
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1276 AA; 140993 MW; 1804D0F011B0FF4E CRC64;

Query Match 36.5%; Score 2453.5; DB 1; Length 1276;
 Best Local Similarity 40.9%; Pred. No. 5.6e-137;
 Matches 520; Conservative 232; Mismatches 461; Indels 57; Gaps 7;
 QY 90 KTOLEIKVNTSFGIMATKMDILIVISTICIAAASFFQRLM----- 136
 Db 23 KKEKKKPAVGVGMFRYADMIDKLCMIITGLAIHIGTLLPLMLVFGMNTDSFTAE 82
 QY 137 -----XOISYDEFYDELTKNLYFYVYLGIEFYVYVYVSTGFIYTGHA 180
 Db 83 ASILPSTNOSGPNSTLIISNSLEEMAVIAYYTGAGVLIYAIQVSLMCLAAQRO 142
 QY 181 TOKIREYTESILRONIGYFPRKLAGEVYTRITADTNLQGISKRGVLTALATPFA 240
 Db 143 IHKIRQKFFHAIMQELIMFVHDVDELNTRITLDVSKINDIGIKIMFQSIYTFAG 202
 QY 241 FIIVAVKKKALICSSITVALVLTLMGGQSOFIYKSKSDYAGGTVAEEVISTRN 300
 Db 203 FLIGISCKLTLVILANSPILGSSALMAKYLSTFKELQAYAKAAVAEEVLAIRT 262
 QY 301 ATAFGTODKLAKQYEVHLDDEAKMGTKNQIYVGMIGAMFGIMVSNYGLGFMGSRFLVD 360

```

Db 263 VIALGGQKLEKRNKMLEKAKVGIKALTASTISIGIALVLYASALAFWGTSLVLS 322
QY 361 GAVVDGILTVLMAIILIGSFLGNVSPNAOFTNAVAAAIFGTIDRQSLDPYSECK 420
Db 323 NEXSIEBVLVFFSILGTSTIGHLANIENAFANARGALEIKIIDNESIDSFISKGY 382
QY 421 TLDFBCHIELRNVKHIYSPREPVYEDVSLMPACKTTALVPGSGSKSTVGLVERF 480
Db 383 KPDSINGNLEFKVHFVPSRSEVOIILKGLNLYKSGQVALVANGSGCKSTVQLMQL 442
QY 481 YMPVRGVLLDGDIDKIDNRLRQOISLVSOBPLYEGTIVYKINIRHGLGTKEENSED 540
Db 443 YDPLEGVASIDGODIRTNVRLREITIGVSOBPLYEPATTIAENIRG-----RED 493
QY 541 KYBELIENAKMANADFTALPEGETVNGQGFLLSGQOKORIAIARAIVSDPKILL 600
Db 494 VTMDELEKAKVKAANAYDFIMKLPHQFDTLVNGEQAOLSGQOKORIAIARAIVNPKILL 553
QY 601 DEATYSLDTRKSEGVQAALERRAEGRTTYIAHRLSTIKTAHNIIVLVNKGIAEGTIDE 660
Db 554 DEATYSLDTRKSEGVQAALERRAEGRTTYIAHRLSTIKTAHNIIVLVNKGIAEGTIDE 613
QY 661 LVDRGAVRYLVEAO--RINQKREAD-----ALBEDAEDLTNADIAKIKTASSSDLDGK 715
Db 614 LMEKCIYELVMTQTRNGNIEPGNNAYGSDPTDASELTSESKSPLIRRSIYRSHRK 673
QY 716 PTYIDRTGTHKVSASAILSRKPETTPKYSLMTLTKFVAFNRPETPYMLIGVSVLAG 775
Db 674 QOQERLSMKEND-----EDVPVLSFWIL-----NLNLSMPYLLVGVCAVING 720
QY 776 GGOPTAVLYAKAISTLSPEOSYSLKRDADWFLMFVVGIIQPTOSTNGAAFAVCS 835
Db 721 CIOPAIYVFSRIYGVFSRDDHETK--RQNCNFSLEFLVAGLISFTYFQGFTEGKAG 779
QY 836 ERLIRBARSAFTTILRODIAFEDKENSNGALTSTLSTETKILSGSVTGLTILMTST 895
Db 780 EILTKVVRVWFVSMKLRQDLSWFDHKNSTGSLTTRIASASSVKAGMRLAVVQNA 839
QY 896 TLGAAIILALATGKMLALYCVVYLLACGFYFYMALQFOSRSKLYAGSANFACAT 955
Db 840 NGTGVILSLVGMQLLVLVLIPIVIGLIEMLKLSQALKKOLEISKINTEAT 899
QY 956 SSTRVYASLTREDDWEIYHAQIDAQGRISLVLSNLSLVSQAALVFECVALGFWYGG 1015
Db 900 ENFRIVSLTRQKFEFTMAQSLQVYRNAMKRAHYGILTFSTQAMTFSYAACRFGA 959
QY 1016 TLIGHHEYDIFRFVCFSELLEGAOSAGYVFSRADMCKAKNAAEFRRLFDKPOIDM 1075
Db 960 YLVAQOLMFPENVMVFSAVFEGAMAGNTSSFAPIYAKAKVASHIRIETPEIDST 1019
QY 1076 SEEGKLETVGEIEFRNHFRTYRPEQVYLRGLDLTVKPGQYVALVPGSGCKSTTIA 1135
Db 1020 STFGKLPITLLEGVKNVNGVFNPTRPNIPLYGLSLFKVKKGTALVVGSGCKSTVQ 1079
QY 1136 LLERFDAIAGSLVNGKDISKININSYFSLVSQOEPPLYOGTKEINTLIGIVEDDV 1195
Db 1080 LLERFDPNAGSYFLDGKEIKQLNVOMLRAHLGLVISOEPLIFDCSIAENIAYDINSRAVS 1139
QY 1196 EEFLIACDANIYDFITMSLPEGFNTVSGKGMLSGGOKORAIARALLRDKILLDE 1255
Db 1140 HEEIYAAKEANIHOFIDSLPKYNTRVGDKQLSGGOKORIAIARALVROPHILLDE 1199
QY 1256 ATVSALDSESEKVVQALDAARGRITIAVAHRLSTIQKADVIVFPGKIVESGISELY 1315
Db 1200 ATVSALDSESEKVVQALDAARGRITIAVAHRLSTIQKADVIVFPGKIVESGISELY 1259
QY 1316 OKKGRYELV 1325
Db 1260 AOKGIFSMV 1269

```

```

MDR2_CRIGR
ID MDR2_CRIGR STANDARD; PRT; 1276 AA.
AC P21449;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 2 (P-glycoprotein 2).
GN PGY2 OR PGY2.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_Taxid=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
   gene family."
RL DNA Seq. 2:89-101(1991).
RN [2]
RP SEQUENCE OF 622-1276 FROM N.A.
RX MEDLINE=88122132; PubMed=2893225;
RA Endicott J.A., Uranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
   Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
   drug-sensitive Chinese hamster ovary cells."
RL Mol. Cell. Biol. 7:4075-4081(1987).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
   DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGY ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
   CAPABILITIES: PGY1 AND PGY2 CAN MEDIATE MDR, WHILE PGY3 APPARENTLY
   CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60041; AAA68884.1; -
DR EMBL; M17896; AAA37007.1; -
DR PIR; B27126; DVHY2C.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transport.
DR InterPro; IPR001140; ABC_transport.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.
DR ProDom; PD000006; ABC_membrane; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
   Multigene family.
FT DOMAIN 1 51
FT TRANSMEM 52 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 409 429 POTENTIAL.
FT TRANSMEM 430 450 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 472 492 POTENTIAL.
FT TRANSMEM 493 513 POTENTIAL.
FT TRANSMEM 514 534 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
FT TRANSMEM 556 576 POTENTIAL.
FT TRANSMEM 577 597 POTENTIAL.
FT TRANSMEM 598 618 POTENTIAL.
FT TRANSMEM 619 639 POTENTIAL.
FT TRANSMEM 640 660 POTENTIAL.
FT TRANSMEM 661 681 POTENTIAL.
FT TRANSMEM 682 702 POTENTIAL.
FT TRANSMEM 703 723 POTENTIAL.
FT TRANSMEM 724 744 POTENTIAL.
FT TRANSMEM 745 765 POTENTIAL.
FT TRANSMEM 766 786 POTENTIAL.
FT TRANSMEM 787 807 POTENTIAL.
FT TRANSMEM 808 828 POTENTIAL.
FT TRANSMEM 829 849 POTENTIAL.
FT TRANSMEM 850 870 POTENTIAL.
FT TRANSMEM 871 891 POTENTIAL.
FT TRANSMEM 892 912 POTENTIAL.
FT TRANSMEM 913 933 POTENTIAL.
FT TRANSMEM 934 954 POTENTIAL.
FT TRANSMEM 955 975 POTENTIAL.
FT TRANSMEM 976 996 POTENTIAL.
FT TRANSMEM 997 1017 POTENTIAL.
FT TRANSMEM 1018 1038 POTENTIAL.
FT TRANSMEM 1039 1059 POTENTIAL.
FT TRANSMEM 1060 1080 POTENTIAL.
FT TRANSMEM 1081 1101 POTENTIAL.
FT TRANSMEM 1102 1122 POTENTIAL.
FT TRANSMEM 1123 1143 POTENTIAL.
FT TRANSMEM 1144 1164 POTENTIAL.
FT TRANSMEM 1165 1185 POTENTIAL.
FT TRANSMEM 1186 1206 POTENTIAL.
FT TRANSMEM 1207 1227 POTENTIAL.
FT TRANSMEM 1228 1248 POTENTIAL.
FT TRANSMEM 1249 1269 POTENTIAL.
FT TRANSMEM 1270 1290 POTENTIAL.
FT TRANSMEM 1291 1311 POTENTIAL.
FT TRANSMEM 1312 1332 POTENTIAL.
FT TRANSMEM 1333 1353 POTENTIAL.
FT TRANSMEM 1354 1374 POTENTIAL.
FT TRANSMEM 1375 1395 POTENTIAL.
FT TRANSMEM 1396 1416 POTENTIAL.
FT TRANSMEM 1417 1437 POTENTIAL.
FT TRANSMEM 1438 1458 POTENTIAL.
FT TRANSMEM 1459 1479 POTENTIAL.
FT TRANSMEM 1480 1500 POTENTIAL.
FT TRANSMEM 1501 1521 POTENTIAL.
FT TRANSMEM 1522 1542 POTENTIAL.
FT TRANSMEM 1543 1563 POTENTIAL.
FT TRANSMEM 1564 1584 POTENTIAL.
FT TRANSMEM 1585 1605 POTENTIAL.
FT TRANSMEM 1606 1626 POTENTIAL.
FT TRANSMEM 1627 1647 POTENTIAL.
FT TRANSMEM 1648 1668 POTENTIAL.
FT TRANSMEM 1669 1689 POTENTIAL.
FT TRANSMEM 1690 1710 POTENTIAL.
FT TRANSMEM 1711 1731 POTENTIAL.
FT TRANSMEM 1732 1752 POTENTIAL.
FT TRANSMEM 1753 1773 POTENTIAL.
FT TRANSMEM 1774 1794 POTENTIAL.
FT TRANSMEM 1795 1815 POTENTIAL.
FT TRANSMEM 1816 1836 POTENTIAL.
FT TRANSMEM 1837 1857 POTENTIAL.
FT TRANSMEM 1858 1878 POTENTIAL.
FT TRANSMEM 1879 1899 POTENTIAL.
FT TRANSMEM 1900 1920 POTENTIAL.
FT TRANSMEM 1921 1941 POTENTIAL.
FT TRANSMEM 1942 1962 POTENTIAL.
FT TRANSMEM 1963 1983 POTENTIAL.
FT TRANSMEM 1984 2004 POTENTIAL.
FT TRANSMEM 2005 2025 POTENTIAL.
FT TRANSMEM 2026 2046 POTENTIAL.
FT TRANSMEM 2047 2067 POTENTIAL.
FT TRANSMEM 2068 2088 POTENTIAL.
FT TRANSMEM 2089 2109 POTENTIAL.
FT TRANSMEM 2110 2130 POTENTIAL.
FT TRANSMEM 2131 2151 POTENTIAL.
FT TRANSMEM 2152 2172 POTENTIAL.
FT TRANSMEM 2173 2193 POTENTIAL.
FT TRANSMEM 2194 2214 POTENTIAL.
FT TRANSMEM 2215 2235 POTENTIAL.
FT TRANSMEM 2236 2256 POTENTIAL.
FT TRANSMEM 2257 2277 POTENTIAL.
FT TRANSMEM 2278 2298 POTENTIAL.
FT TRANSMEM 2299 2319 POTENTIAL.
FT TRANSMEM 2320 2340 POTENTIAL.
FT TRANSMEM 2341 2361 POTENTIAL.
FT TRANSMEM 2362 2382 POTENTIAL.
FT TRANSMEM 2383 2403 POTENTIAL.
FT TRANSMEM 2404 2424 POTENTIAL.
FT TRANSMEM 2425 2445 POTENTIAL.
FT TRANSMEM 2446 2466 POTENTIAL.
FT TRANSMEM 2467 2487 POTENTIAL.
FT TRANSMEM 2488 2508 POTENTIAL.
FT TRANSMEM 2509 2529 POTENTIAL.
FT TRANSMEM 2530 2550 POTENTIAL.
FT TRANSMEM 2551 2571 POTENTIAL.
FT TRANSMEM 2572 2592 POTENTIAL.
FT TRANSMEM 2593 2613 POTENTIAL.
FT TRANSMEM 2614 2634 POTENTIAL.
FT TRANSMEM 2635 2655 POTENTIAL.
FT TRANSMEM 2656 2676 POTENTIAL.
FT TRANSMEM 2677 2697 POTENTIAL.
FT TRANSMEM 2698 2718 POTENTIAL.
FT TRANSMEM 2719 2739 POTENTIAL.
FT TRANSMEM 2740 2760 POTENTIAL.
FT TRANSMEM 2761 2781 POTENTIAL.
FT TRANSMEM 2782 2802 POTENTIAL.
FT TRANSMEM 2803 2823 POTENTIAL.
FT TRANSMEM 2824 2844 POTENTIAL.
FT TRANSMEM 2845 2865 POTENTIAL.
FT TRANSMEM 2866 2886 POTENTIAL.
FT TRANSMEM 2887 2907 POTENTIAL.
FT TRANSMEM 2908 2928 POTENTIAL.
FT TRANSMEM 2929 2949 POTENTIAL.
FT TRANSMEM 2950 2970 POTENTIAL.
FT TRANSMEM 2971 2991 POTENTIAL.
FT TRANSMEM 2992 3012 POTENTIAL.
FT TRANSMEM 3013 3033 POTENTIAL.
FT TRANSMEM 3034 3054 POTENTIAL.
FT TRANSMEM 3055 3075 POTENTIAL.
FT TRANSMEM 3076 3096 POTENTIAL.
FT TRANSMEM 3097 3117 POTENTIAL.
FT TRANSMEM 3118 3138 POTENTIAL.
FT TRANSMEM 3139 3159 POTENTIAL.
FT TRANSMEM 3160 3180 POTENTIAL.
FT TRANSMEM 3181 3201 POTENTIAL.
FT TRANSMEM 3202 3222 POTENTIAL.
FT TRANSMEM 3223 3243 POTENTIAL.
FT TRANSMEM 3244 3264 POTENTIAL.
FT TRANSMEM 3265 3285 POTENTIAL.
FT TRANSMEM 3286 3306 POTENTIAL.
FT TRANSMEM 3307 3327 POTENTIAL.
FT TRANSMEM 3328 3348 POTENTIAL.
FT TRANSMEM 3349 3369 POTENTIAL.
FT TRANSMEM 3370 3390 POTENTIAL.
FT TRANSMEM 3391 3411 POTENTIAL.
FT TRANSMEM 3412 3432 POTENTIAL.
FT TRANSMEM 3433 3453 POTENTIAL.
FT TRANSMEM 3454 3474 POTENTIAL.
FT TRANSMEM 3475 3495 POTENTIAL.
FT TRANSMEM 3496 3516 POTENTIAL.
FT TRANSMEM 3517 3537 POTENTIAL.
FT TRANSMEM 3538 3558 POTENTIAL.
FT TRANSMEM 3559 3579 POTENTIAL.
FT TRANSMEM 3580 3600 POTENTIAL.
FT TRANSMEM 3601 3621 POTENTIAL.
FT TRANSMEM 3622 3642 POTENTIAL.
FT TRANSMEM 3643 3663 POTENTIAL.
FT TRANSMEM 3664 3684 POTENTIAL.
FT TRANSMEM 3685 3705 POTENTIAL.
FT TRANSMEM 3706 3726 POTENTIAL.
FT TRANSMEM 3727 3747 POTENTIAL.
FT TRANSMEM 3748 3768 POTENTIAL.
FT TRANSMEM 3769 3789 POTENTIAL.
FT TRANSMEM 3790 3810 POTENTIAL.
FT TRANSMEM 3811 3831 POTENTIAL.
FT TRANSMEM 3832 3852 POTENTIAL.
FT TRANSMEM 3853 3873 POTENTIAL.
FT TRANSMEM 3874 3894 POTENTIAL.
FT TRANSMEM 3895 3915 POTENTIAL.
FT TRANSMEM 3916 3936 POTENTIAL.
FT TRANSMEM 3937 3957 POTENTIAL.
FT TRANSMEM 3958 3978 POTENTIAL.
FT TRANSMEM 3979 3999 POTENTIAL.
FT TRANSMEM 4000 4020 POTENTIAL.
FT TRANSMEM 4021 4041 POTENTIAL.
FT TRANSMEM 4042 4062 POTENTIAL.
FT TRANSMEM 4063 4083 POTENTIAL.
FT TRANSMEM 4084 4104 POTENTIAL.
FT TRANSMEM 4105 4125 POTENTIAL.
FT TRANSMEM 4126 4146 POTENTIAL.
FT TRANSMEM 4147 4167 POTENTIAL.
FT TRANSMEM 4168 4188 POTENTIAL.
FT TRANSMEM 4189 4209 POTENTIAL.
FT TRANSMEM 4210 4230 POTENTIAL.
FT TRANSMEM 4231 4251 POTENTIAL.
FT TRANSMEM 4252 4272 POTENTIAL.
FT TRANSMEM 4273 4293 POTENTIAL.
FT TRANSMEM 4294 4314 POTENTIAL.
FT TRANSMEM 4315 4335 POTENTIAL.
FT TRANSMEM 4336 4356 POTENTIAL.
FT TRANSMEM 4357 4377 POTENTIAL.
FT TRANSMEM 4378 4398 POTENTIAL.
FT TRANSMEM 4399 4419 POTENTIAL.
FT TRANSMEM 4420 4440 POTENTIAL.
FT TRANSMEM 4441 4461 POTENTIAL.
FT TRANSMEM 4462 4482 POTENTIAL.
FT TRANSMEM 4483 4503 POTENTIAL.
FT TRANSMEM 4504 4524 POTENTIAL.
FT TRANSMEM 4525 4545 POTENTIAL.
FT TRANSMEM 4546 4566 POTENTIAL.
FT TRANSMEM 4567 4587 POTENTIAL.
FT TRANSMEM 4588 4608 POTENTIAL.
FT TRANSMEM 4609 4629 POTENTIAL.
FT TRANSMEM 4630 4650 POTENTIAL.
FT TRANSMEM 4651 4671 POTENTIAL.
FT TRANSMEM 4672 4692 POTENTIAL.
FT TRANSMEM 4693 4713 POTENTIAL.
FT TRANSMEM 4714 4734 POTENTIAL.
FT TRANSMEM 4735 4755 POTENTIAL.
FT TRANSMEM 4756 4776 POTENTIAL.
FT TRANSMEM 4777 4797 POTENTIAL.
FT TRANSMEM 4798 4818 POTENTIAL.
FT TRANSMEM 4819 4839 POTENTIAL.
FT TRANSMEM 4840 4860 POTENTIAL.
FT TRANSMEM 4861 4881 POTENTIAL.
FT TRANSMEM 4882 4902 POTENTIAL.
FT TRANSMEM 4903 4923 POTENTIAL.
FT TRANSMEM 4924 4944 POTENTIAL.
FT TRANSMEM 4945 4965 POTENTIAL.
FT TRANSMEM 4966 4986 POTENTIAL.
FT TRANSMEM 4987 5007 POTENTIAL.
FT TRANSMEM 5008 5028 POTENTIAL.
FT TRANSMEM 5029 5049 POTENTIAL.
FT TRANSMEM 5050 5070 POTENTIAL.
FT TRANSMEM 5071 5091 POTENTIAL.
FT TRANSMEM 5092 5112 POTENTIAL.
FT TRANSMEM 5113 5133 POTENTIAL.
FT TRANSMEM 5134 5154 POTENTIAL.
FT TRANSMEM 5155 5175 POTENTIAL.
FT TRANSMEM 5176 5196 POTENTIAL.
FT TRANSMEM 5197 5217 POTENTIAL.
FT TRANSMEM 5218 5238 POTENTIAL.
FT TRANSMEM 5239 5259 POTENTIAL.
FT TRANSMEM 5260 5280 POTENTIAL.
FT TRANSMEM 5281 5301 POTENTIAL.
FT TRANSMEM 5302 5322 POTENTIAL.
FT TRANSMEM 5323 5343 POTENTIAL.
FT TRANSMEM 5344 5364 POTENTIAL.
FT TRANSMEM 5365 5385 POTENTIAL.
FT TRANSMEM 5386 5406 POTENTIAL.
FT TRANSMEM 5407 5427 POTENTIAL.
FT TRANSMEM 5428 5448 POTENTIAL.
FT TRANSMEM 5449 5469 POTENTIAL.
FT TRANSMEM 5470 5490 POTENTIAL.
FT TRANSMEM 5491 5511 POTENTIAL.
FT TRANSMEM 5512 5532 POTENTIAL.
FT TRANSMEM 5533 5553 POTENTIAL.
FT TRANSMEM 5554 5574 POTENTIAL.
FT TRANSMEM 5575 5595 POTENTIAL.
FT TRANSMEM 5596 5616 POTENTIAL.
FT TRANSMEM 5617 5637 POTENTIAL.
FT TRANSMEM 5638 5658 POTENTIAL.
FT TRANSMEM 5659 5679 POTENTIAL.
FT TRANSMEM 5680 5700 POTENTIAL.
FT TRANSMEM 5701 5721 POTENTIAL.
FT TRANSMEM 5722 5742 POTENTIAL.
FT TRANSMEM 5743 5763 POTENTIAL.
FT TRANSMEM 5764 5784 POTENTIAL.
FT TRANSMEM 5785 5805 POTENTIAL.
FT TRANSMEM 5806 5826 POTENTIAL.
FT TRANSMEM 5827 5847 POTENTIAL.
FT TRANSMEM 5848 5868 POTENTIAL.
FT TRANSMEM 5869 5889 POTENTIAL.
FT TRANSMEM 5890 5910 POTENTIAL.
FT TRANSMEM 5911 5931 POTENTIAL.
FT TRANSMEM 5932 5952 POTENTIAL.
FT TRANSMEM 5953 5973 POTENTIAL.
FT TRANSMEM 5974 5994 POTENTIAL.
FT TRANSMEM 5995 6015 POTENTIAL.
FT TRANSMEM 6016 6036 POTENTIAL.
FT TRANSMEM 6037 6057 POTENTIAL.
FT TRANSMEM 6058 6078 POTENTIAL.
FT TRANSMEM 6079 6099 POTENTIAL.
FT TRANSMEM 6100 6120 POTENTIAL.
FT TRANSMEM 6121 6141 POTENTIAL.
FT TRANSMEM 6142 6162 POTENTIAL.
FT TRANSMEM 6163 6183 POTENTIAL.
FT TRANSMEM 6184 6204 POTENTIAL.
FT TRANSMEM 6205 6225 POTENTIAL.
FT TRANSMEM 6226 6246 POTENTIAL.
FT TRANSMEM 6247 6267 POTENTIAL.
FT TRANSMEM 6268 6288 POTENTIAL.
FT TRANSMEM 6289 6309 POTENTIAL.
FT TRANSMEM 6310 6330 POTENTIAL.
FT TRANSMEM 6331 6351 POTENTIAL.
FT TRANSMEM 6352 6372 POTENTIAL.
FT TRANSMEM 6373 6393 POTENTIAL.
FT TRANSMEM 6394 6414 POTENTIAL.
FT TRANSMEM 6415 6435 POTENTIAL.
FT TRANSMEM 6436 6456 POTENTIAL.
FT TRANSMEM 6457 6477 POTENTIAL.
FT TRANSMEM 6478 6498 POTENTIAL.
FT TRANSMEM 6499 6519 POTENTIAL.
FT TRANSMEM 6520 6540 POTENTIAL.
FT TRANSMEM 6541 6561 POTENTIAL.
FT TRANSMEM 6562 6582 POTENTIAL.
FT TRANSMEM 6583 6603 POTENTIAL.
FT TRANSMEM 6604 6624 POTENTIAL.
FT TRANSMEM 6625 6645 POTENTIAL.
FT TRANSMEM 6646 6666 POTENTIAL.
FT TRANSMEM 6667 6687 POTENTIAL.
FT TRANSMEM 6688 6708 POTENTIAL.
FT TRANSMEM 6709 6729 POTENTIAL.
FT TRANSMEM 6730 6750 POTENTIAL.
FT TRANSMEM 6751 6771 POTENTIAL.
FT TRANSMEM 6772 6792 POTENTIAL.
FT TRANSMEM 6793 6813 POTENTIAL.
FT TRANSMEM 6814 6834 POTENTIAL.
FT TRANSMEM 6835 6855 POTENTIAL.
FT TRANSMEM 6856 6876 POTENTIAL.
FT TRANSMEM 6877 6897 POTENTIAL.
FT TRANSMEM 6898 6918 POTENTIAL.
FT TRANSMEM 6919 6939 POTENTIAL.
FT TRANSMEM 6940 6960 POTENTIAL.
FT TRANSMEM 6961 6981 POTENTIAL.
FT TRANSMEM 6982 7002 POTENTIAL.
FT TRANSMEM 7003 7023 POTENTIAL.
FT TRANSMEM 7024 7044 POTENTIAL.
FT TRANSMEM 7045 7065 POTENTIAL.
FT TRANSMEM 7066 7086 POTENTIAL.
FT TRANSMEM 7087 7107 POTENTIAL.
FT TRANSMEM 7108 7128 POTENTIAL.
FT TRANSMEM 7129 7149 POTENTIAL.
FT TRANSMEM 7150 7170 POTENTIAL.
FT TRANSMEM 7171 7191 POTENTIAL.
FT TRANSMEM 7192 7212 POTENTIAL.
FT TRANSMEM 7213 7233 POTENTIAL.
FT TRANSMEM 7234 7254 POTENTIAL.
FT TRANSMEM 7255 7275 POTENTIAL.
FT TRANSMEM 7276 7296 POTENTIAL.
FT TRANSMEM 7297 7317 POTENTIAL.
FT TRANSMEM 7318 7338 POTENTIAL.
FT TRANSMEM 7339 7359 POTENTIAL.
FT TRANSMEM 7360 7380 POTENTIAL.
FT TRANSMEM 7381 7401 POTENTIAL.
FT TRANSMEM 7402 7422 POTENTIAL.
FT TRANSMEM 7423 7443 POTENTIAL.
FT TRANSMEM 7444 7464 POTENTIAL.
FT TRANSMEM 7465 7485 POTENTIAL.
FT TRANSMEM 7486 7506 POTENTIAL.
FT TRANSMEM 7507 7527 POTENTIAL.
FT TRANSMEM 7528 7548 POTENTIAL.
FT TRANSMEM 7549 7569 POTENTIAL.
FT TRANSMEM 7570 7590 POTENTIAL.
FT TRANSMEM 7591 7611 POTENTIAL.
FT TRANSMEM 7612 7632 POTENTIAL.
FT TRANSMEM 7633 7653 POTENTIAL.
FT TRANSMEM 7654 7674 POTENTIAL.
FT TRANSMEM 7675 7695 POTENTIAL.
FT TRANSMEM 7696 7716 POTENTIAL.
FT TRANSMEM 7717 7737 POTENTIAL.
FT TRANSMEM 7738 7758 POTENTIAL.
FT TRANSMEM 7759 7779 POTENTIAL.
FT TRANSMEM 7780 7800 POTENTIAL.
FT TRANSMEM 7801 7821 POTENTIAL.
FT TRANSMEM 7822 7842 POTENTIAL.
FT TRANSMEM 7843 7863 POTENTIAL.
FT TRANSMEM 7864 7884 POTENTIAL.
FT TRANSMEM 7885 7905 POTENTIAL.
FT TRANSMEM 7906 7926 POTENTIAL.
FT TRANSMEM 7927 7947 POTENTIAL.
FT TRANSMEM 7948 7968 POTENTIAL.
FT TRANSMEM 7969 7989 POTENTIAL.
FT TRANSMEM 7990 8010 POTENTIAL.
FT TRANSMEM 8011 8031 POTENTIAL.
FT TRANSMEM 8032 8052 POTENTIAL.
FT TRANSMEM 8053 8073 POTENTIAL.
FT TRANSMEM 8074 8094 POTENTIAL.
FT TRANSMEM 8095 8115 POTENTIAL.
FT TRANSMEM 8116 8136 POTENTIAL.
FT TRANSMEM 8137 8157 POTENTIAL.
FT TRANSMEM 8158 8178 POTENTIAL.
FT TRANSMEM 8179 8199 POTENTIAL.
FT TRANSMEM 8200 8220 POTENTIAL.
FT TRANSMEM 8221 8241 POTENTIAL.
FT TRANSMEM 8242 8262 POTENTIAL.
FT TRANSMEM 8263 8283 POTENTIAL.
FT TRANSMEM 8284 8304 POTENTIAL.
FT TRANSMEM 8305 8325 POTENTIAL.
FT TRANSMEM 8326 8346 POTENTIAL.
FT TRANSMEM 8347 8367 POTENTIAL.
FT TRANSMEM 8368 8388 POTENTIAL.
FT TRANSMEM 8389 8409 POTENTIAL.
FT TRANSMEM 8410 8430 POTENTIAL.
FT TRANSMEM 8431 8451 POTENTIAL.
FT TRANSMEM 8452 8472 POTENTIAL.
FT TRANSMEM 8473 8493 POTENTIAL.
FT TRANSMEM 8494 8514 POTENTIAL.
FT TRANSMEM 8515 8535 POTENTIAL.
FT TRANSMEM 8536 8556 POTENTIAL.
FT TRANSMEM 8557 8577 POTENTIAL.
FT TRANSMEM 8578 8598 POTENTIAL.
FT TRANSMEM 8599 8619 POTENTIAL.
FT TRANSMEM 8620 8640 POTENTIAL.
FT TRANSMEM 8641 8661 POTENTIAL.
FT TRANSMEM 8662 8682 POTENTIAL.
FT TRANSMEM 8683 8703 POTENTIAL.
FT TRANSMEM 8704 8724 POTENTIAL.
FT TRANSMEM 8725 8745 POTENTIAL.
FT TRANSMEM 8746 8766 POTENTIAL.
FT TRANSMEM 8767 8787 POTENTIAL.
FT TRANSMEM 8788 8808 POTENTIAL.
FT TRANSMEM 8809 8829 POTENTIAL.
FT TRANSMEM 8830 8850 POTENTIAL.
FT TRANSMEM 8851 8871 POTENTIAL.
FT TRANSMEM 8872 8892 POTENTIAL.
FT TRANSMEM 8893 8913 POTENTIAL.
FT TRANSMEM 8914 8934 POTENTIAL.
FT TRANSMEM 8935 8955 POTENTIAL.
FT TRANSMEM 8956 8976 POTENTIAL.
FT TRANSMEM 8977 8997 POTENTIAL.
FT TRANSMEM 8998 9018 POTENTIAL.
FT TRANSMEM 9019 9039 POTENTIAL.
FT TRANSMEM 9040 9060 POTENTIAL.
FT TRANSMEM 9061 9081 POTENTIAL.
FT TRANSMEM 9082 9102 POTENTIAL.
FT TRANSMEM 9103 9123 POTENTIAL.
FT TRANSMEM 9124 9144 POTENTIAL.
FT TRANSMEM 9145 9165 POTENTIAL.
FT TRANSMEM 9166 9186 POTENTIAL.
FT TRANSMEM 9187 9207 POTENTIAL.
FT TRANSMEM 9208 9228 POTENTIAL.
FT TRANSMEM 9229 9249 POTENTIAL.
FT TRANSMEM 9250 9270 POTENTIAL.
FT TRANSMEM 9271 9291 POTENTIAL.
FT TRANSMEM 9292 9312 POTENTIAL.
FT TRANSMEM 9313 9333 POTENTIAL.
FT TRANSMEM 9334 9354 POTENTIAL.
FT TRANSMEM 9355 9375 POTENTIAL.
FT TRANSMEM 9376 9396 POTENTIAL.
FT TRANSMEM 9397 9417 POTENTIAL.
FT TRANSMEM 9418 9438 POTENTIAL.
FT TRANSMEM 9439 9459 POTENTIAL.
FT TRANSMEM 9460 9480 POTENTIAL.
FT TRANSMEM 9481 9501 POTENTIAL.
FT TRANSMEM 9502 9522 POTENTIAL.
FT TRANSMEM 9523 9543 POTENTIAL.
FT TRANSMEM 9544 9564 POTENTIAL.
FT TRANSMEM 9565 9585 POTENTIAL.
FT TRANSMEM 9586 9606 POTENTIAL.
FT TRANSMEM 9607 9627 POTENTIAL.
FT TRANSMEM 9628 9648 POTENTIAL.
FT TRANSMEM 9649 9669 POTENTIAL.
FT TRANSMEM 9670 9690 POTENTIAL.
FT TRANSMEM 9691 9711 POTENTIAL.
FT TRANSMEM 9712 9732 POTENTIAL.
FT TRANSMEM 9733 9753 POTENTIAL.
FT TRANSMEM 9754 9774 POTENTIAL.
FT TRANSMEM 9775 9795 POTENTIAL.
FT TRANSMEM 9796 9816 POTENTIAL.
FT TRANSMEM 9817 9837 POTENTIAL.
FT TRANSMEM 9838 9858 POTENTIAL.
FT TRANSMEM 9859 9879 POTENTIAL.
FT TRANSMEM 9880 9900 POTENTIAL.
FT TRANSMEM 9901 9921 POTENTIAL.
FT TRANSMEM 9922 9942 POTENTIAL.
FT TRANSMEM 9943 9963 POTENTIAL.
FT TRANSMEM 9964 9984 POTENTIAL.
FT TRANSMEM 9985 10000 POTENTIAL.

```

RESULT 6


```

FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 71 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT DOMAIN 348 709 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 710 730 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 832 852 POTENTIAL.
FT TRANSMEM 854 874 POTENTIAL.
FT TRANSMEM 937 957 POTENTIAL.
FT TRANSMEM 968 988 POTENTIAL.
FT DOMAIN 989 1277 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (BY SIMILARITY).
FT NP_BIND 1070 1077 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1277
FT CAROHD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1277 AA; 141386 MW; 8AFDD619D2934C1 CRC64;

Query Match 35.6%; Score 2392.5; DB 1; Length 1277;
Best Local Similarity 40.4%; Pred. No. 2,3e-133;
Matches 513; Conservative 236; Mismatches 459; Indels 61; Gaps 10;

QY 94 EEIKVNISEFGIMRWYATKMDILMISTICATAAASFQIRIML----- 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 26 KEKRPVAGIFGMRYADMIDKLCMALGTALAIHGLTLPMLFQYMDSPFSDPHS 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 -----YQSYDEFDDELTKNLYFPYLGIGEFVYVYVSVGFIYGEHATQK 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 86 DRAITNQEINSTRHVSDFSLIEDMAMVAYYYIGAGVLAIVAYIOVSLMCLAAGROIHR 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 IREYVLESLIRONIGYFDKLAGEVTRITADNLIODGISEKVGTLTALATFYAFII 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 146 IROKFEHALMNGEIGFVDNAGELNTRLDVSKINDGIDKLGFMFPOSITTFSGFTI 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 AYYKYKLLALICSTVALVLTWGGSOPTIKYSKSLDSYGAGTVAEVVISINATA 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 206 GFISGKRLVLIVLAPSLIGLSSAMAKVLTSTPKELQAYAKAGVAEVEVLAIRTVIA 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 FGTQDOLAOYEVLHDEAKMGWTKNOIVGMFGAMFGMYSGVGFMMGSRFLVDGAV 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 266 FGGQKKELEKYNKLEAKRVGKIKATIANISIGIAYLVATYALAFMVGISLVISNEY 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 364 DWGDLITVLMALIGSFSGVSNVSPNAQAFITNNAVAAAKIFGTIDROSPIDPNSNEKTI 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 326 SIGOVLTFEFSILGTFISGLHAPNIEAFANAGAYEIFKIIDNPSIDSEFTKHKRP 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 424 HFECHLELNKHIYSPREPVYVMEVYSLSMPAKGTTALVGSFGSKTVYGLVERFYMP 483
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 386 SIMGNLEFKNVYFNYSREBEVKLNLKYNKSGQTVAVLVGNSGCKSTVOQLQRLYDP 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 484 VSGTVLLDGHDKDLNMLRQOISLVSOEPVLFGTITKYNHGLIGIKYENESDKYR 543
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 446 IEGEVSIDODIKTIVKRLREITIGVSOEPVLFATTAENITYG-----RENVYMD 498
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 544 ELIENAKANANADFTTALPEGEYEVNGVGFELSGGOKORAIARAVYSDPKILLDDA 603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 499 --IEKAVKEANAYDFIMKLPKFDLVEGEGAGLSSGGOKORAIARAVYSDPKILLDDA 566
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 604 TSAIDTKSGCVQOALERAEGFTTVIAHRLSTIKTANIVYLVNGKTAEGSTHDELYD 663
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 557 TSAIDTESAVVOALDKAREGFTTVIAHRLSTVNAVIVGDFGDIVEQGNHEELMK 616
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 664 RGAAYRKLVEAQ-RINE-----QKADALDADADEDLTNDIAKIKTASASSDLDKPT 718
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 617 EKGITFKLWOTRGNEIIPGNNAVYSDTQASELTSESSSPLIRISIRSIHRQDQ 676
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 719 IDRTGTHKSVSSALSKRPETTPKYSLTLLKTFVASFNRPEIPIYMLIGLVEVLGAGQ 778
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 677 ERLLSKEVD-----EDVPVMSFWQILK-----LNISEPVLVGVLCVINGCIG 723

```

```

QY 779 PTOGAVLYAKAISTLSPESQYSLRHADFWSLMFEVVGIIQFIOTSTNGAFAWCSERL 838
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 724 PYVAIYFSLVGVFSRDDHETKQR-NCNLFSLFLVMGMISVYTFPGQFFPKAGGELL 782
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 839 IRRARSTARITIIRODIAFDEKENSTGALTSFTSTYHLSGVSVTLGITLMSTYLG 898
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 783 TKRLRMRVMSMLRODISFDDHKNTGSLTRLASDASNVMGMSRLAVYQYVANNG 842
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 899 AATIIALAI--GWLKALVGVSVVPLVACGEYFVYLAQFQSKRLAYGSANFACENS 956
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 843 TGLILSLVLYQWGLLVITPIVLVGLIEMKLSQALKKKELEISGKIATEALE 902
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 957 SIRTVASLTRERDWEIYHQAQDQRTSLISYRSLLYASSQALVFECVALGFYWGCT 1016
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 903 NERTVAVSLREQEFTETVMAQSIIQIPYRNALMKRAHVGIFAFATQAMIVSYAACFRGAY 962
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1017 LGHHEYDIFRFVGVCSSEILFGAQSAGYFSPAPDMGKAKNAAEFRRLFDKRPQIDMNS 1076
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 963 LVARELMEFENWLVSAVAFGMAAGNTSFPADYAKAKVSAHSHTIGIIEKIPEDSYS 1022
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1077 EEEGKLEFVEGEIEFRNVRHRYPTREOPVLRGLDLTVRKGOVALVGPSSGCKSTTIAL 1136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1023 TEGLRKMWLEGNKFKFNGVAFNPTRPNIPIVGLSFEVAKKGTILRVGSSGCKSTVQVL 1082
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1137 LERFYDALAGSILVDKDISKLINSYRSFLSVSQEPITLYOGTIKENILGIVEDDVE 1196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1083 LERFYDPMAGTVFLDGEKEIKQLNVQCVRA-LGIVSQEPITLPCDSIENIAYGDSRVSH 1141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1197 EPLIKACKDANIYDFIMSLPEGFNTYVSGKGMGLSGGOKORAIARALLRDKRIILLDDA 1256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1142 EETVRARENNHQFIDSLPEKYNTRVGDQGLSGGOKORAIARALLRQPIILLDDA 1201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1257 TSAIDSESEKVVQOALDAARGTITIAVHRLSTIKAVIYVFDQKIVESGTHSELVQ 1316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1202 TSAIDTESKRYOGEALDKAREGRTCVIAHRLSTIQNADLIYIQNGVKEHSTHQQLLA 1261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1317 KKGRIEYLV 1325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1262 OKGIYFSMV 1270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
MDR2_MOUSE
ID MDR2_MOUSE STANDARD; PRT; 1276 AA.
AC P21440;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 2 (P-glycoprotein 2).
GN ABCB4 OR PGY2 OR PGY-2 OR MDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_Taxid:10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88302195; PubMed=3405218;
RT Gros P., Raymond M., Bell J., Housman D.;
RT "Cloning and characterization of a second member of the mouse mdr
RT gene family.";
RL Mol. Cell. Biol. 8:2770-2778(1988).
RN [2]
RP SEQUENCE OF 1-23 FROM N.A.
RA STRAIN=BA16/c;
RA Kirschner L.S., Horwitz S.B.;
RA Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.
CC -I- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS
CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.

```


CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 DR EMBL: J03398; AAA39516.1; -
 DR EMBL: M74151; AAA39515.1; -
 DR PIR: A30409; DVMS2.
 DR HSSP: P13569; INBD.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family.
 FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 52 75 POTENTIAL.
 FT TRANSMEM 119 139 POTENTIAL.
 FT TRANSMEM 188 208 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 296 316 POTENTIAL.
 FT TRANSMEM 325 345 POTENTIAL.
 FT DOMAIN 346 707 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 708 728 POTENTIAL.
 FT TRANSMEM 753 773 POTENTIAL.
 FT TRANSMEM 829 849 POTENTIAL.
 FT TRANSMEM 850 870 POTENTIAL.
 FT TRANSMEM 933 953 POTENTIAL.
 FT TRANSMEM 970 990 POTENTIAL.
 FT DOMAIN 991 1276 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 426 433 ATP (BY SIMILARITY).
 FT NP_BIND 1066 1073 ATP (BY SIMILARITY).
 FT REPEAT 636 1276
 FT SEQUENCE 1276 AA; 140332 MW; A6C38DCD2C118EEF CRC64;
 Query Match 35.58; Score 2387; DB 1; Length 1276;
 Best Local Similarity 42.18; Pred. No. 4.8e-133;
 Matches 534; Conservative 212; Mismatches 458; Indels 64; Gaps 13;
 QY KVN-ISPFGIMRYATKMDILIMVISTICATAAATFORIML-VQISYDEYD----- 146
 DB 33 KVLILGLITLFRISDMODKLFMLGLITLMAIHAHSGPLMLMIVGEMTDKVTNDNGNSLP 92
 QY 147 -----ELTRNLVEYVYLGIEFTVYVSTGVETGTEHATOKIREYYLE 190
 DB 93 VNFSLMLNGLRIIEEMETRYAYVYSGLGCVLYAAYIQVSEFWTLAAGROIKIKRQKFPH 152
 QY 191 SIIRONGYDCKLGAGVTRITADTDLIODGISEKGLTLTALATVTVTAIIVYVYWK 250
 DB 153 AIIROEGWMDIRGTETLNRITDVSRISEIGIDKGMFEFOAIATFPAFIVGFIGWK 212
 QY 251 IALICSTYVALVLMGGSOFIIRYKSKISDYGAGTAAEEVYISSIRNATVAGTQDKL 310
 DB 213 LTVIYMAISPIIGISTVYVAKKILSTFSDKELAAVAKGAAVEEPGAIKRVIAFGGONKE 272
 QY 311 AKQYEVHLEAEKWKGTQIVGMFGIMGFLMYSNYGLGFMGSRPLVDGADVGDILT 370
 DB 273 LERQKHLERNAKIKIGIKKAIISANISMGIAFLIYASVALAFWGSTLVISKEYVIGNAMT 332
 QY 371 VLAIIIGSSISGVSNAPNAFTNAVAARIGTIDROSLDIPYISNEGKTLDFEGSHIE 430
 DB 333 VFSSILIGASVGOAACPIDAFANARGAAYVIEDIDNNPKIDFSERGHKPDNIKGNLE 392

QY 431 LKNVKKIYPSRPEYTVMEDVSLSPAGKTALVSPSSGSKTVGLYERFMPYRGTVLL 490
 DB 393 FSDVHFETPSRANKIKILKLNIAKYSQGYALVNSGCGSTTVQILQRLDIPREGIST 452
 QY 491 DGHDKDLNRWMLRQOISLVSQEPVLEGTITTKYKIRHGLIGTKYNESEDKVRELIENAA 550
 DB 453 DGDIRNFNRCLREIIGVVSQEPVLESTTIAENIRFG-----RGVNTMDE----EKAV 503
 QY 551 KMANHDFITALPBGTEINVGORFLLSGGOKQALAIARAVSPKILLDEATSAIDTK 610
 DB 504 KEANAYDFIMKLPGKFPTLVGDRCQAOLSGGOKRIAIARLVNPKILLDEATSAIDTE 563
 QY 611 SEGVQAALBRAAGRTTIIAHRSTIKTAHNIVLVNKGIAOGTHDELVDNGAYRK 670
 DB 564 SEAEVQAALDKARGRRTTIIAHRSTIRNADYIAGEEDGYIVQSGHSELMKKEGIYFR 623
 QY 671 LVEAQRINEQKEDALDADAEGLTNADIAKIKTAYSSASSDLDGKPTTIDRTGHSVSS 730
 DB 624 LVNQTAGSQILSEFEVELSDEKAGDVAP-----NGMKARIFRNSTKSKLS 672
 QY 731 ALLSKRPER-----TPKSLMTLLKFVASFNRPETPLMILIGVEVYLAGGQPTQAVL 784
 DB 673 PHQRILDEETNELDANVPVPSFLKVLK-----LNKTEPYPVGVCAIANGALOPAFSII 728
 QY 785 YAKAISTLSIPESQSKLRHDADPMSIMFVYVGIQFTOSTNGAFAVCSERLIRARS 844
 DB 729 LSEMLAIFG-PGDDAVK-QQKKNFSLVFLGLVLSFTFPLOGFTGKAGEIITLRLS 786
 QY 845 TAFRTILRODIAEFDEKENSTGALTSLSTETKHLSGVSGVLTGTLTMTSTTIGAAITIA 904
 DB 787 MAFKAMLRQMSWPDHDKNSTGALTSLRADAQVQATGATKLLAIQNTANLGTGIIS 846
 QY 905 LAIGMKLALVCISVVPILACGFRTFMALQFQSRKLAIEGSAFNFCATSTRYASL 964
 DB 847 FIYGMQTLTLLSVVPPIAAGIVEMKMLAGNAKROKKEMAAKIKTAEIENRTYVSL 906
 QY 965 TREHDVEIYHAQDAGRTSLISLRSLLYASSQALVFECVALGFMYGGTLL--GHHE 1022
 DB 907 TQEKKEFSMYEKHLGGRYSVRKAHAYIGTIFSQAFMFTSYAGCTRFSGSLIVNCHMR 966
 QY 1023 Y-DIREFVCFSEILFQAQSGATVSEFAPDMGKAKNAAEFRRLDERRKPOIDNMSEGEK 1081
 DB 967 FKDY---ILVFSAIVLGAVALGHASSFAPDPAKAKLSAAYLFSLEHQPLIDSYSGGLW 1023
 QY 1082 LQYEGEIEERNNHFRPTPEQVYVGLDLTVKPGQYVALVSGCGKSTTIALLERFY 1141
 DB 1024 PDKEGSGVTENEVFNPTIRANVPVLOGLSLEVKKGQTLALVSGCGKSTVQOLLERFY 1083
 QY 1142 DALGSLIVGKDISKININSYRSFLSVSOEPTLYGCTIKENTLLGIVEDDVEEFLIK 1201
 DB 1084 DPMAGSVLLDGOEAKKILANQMLQGLIVSOEPLIFPCSTIAENIAYDNGSRVPHDITVR 1143
 QY 1202 ACKDANTIDYIMSLPGEFNTVYSGKGMGSGGOKORVAIRARALLRDKPILLDEATSAID 1261
 DB 1144 AAKRANIHPIETLTPQYKTVRVGDKGTQLSGGOKRIAIARALRROPVILLDEATSAID 1203
 QY 1262 SESSEKVVQAALDAARRTTIIAHRSTIQKADYIVYPOGKIVESGTSSELVQKGRY 1321
 DB 1204 TESEKVVQEAALDKARBRGTCIVIAHRLSTIONMDLIVIEGKVKKEGTQOOLLAQIGY 1263
 QY 1322 YELVNIQS 1329
 DB 1264 FSNWNIQA 1271
 RESULT 9
 MDR3_HUMAN
 ID MDR3_HUMAN STANDARD; PRT; 1279 AA.
 AC P21439;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Multidrug resistance protein 3 (P-glycoprotein 3).
GN ABCB4 OR PGY3 OR MDR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; PubMed=2906314;
RA van der Bliek A.M., Koelman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL Gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=8811519; PubMed=2892668;
RA van der Bliek A.M., Baas F., ten Houte de Lange T., Koelman P.M.,
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL EMBO J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; PubMed=2002063;
RA Lincke C.R., Smit J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT
CC CAPABLE OF CONFERRING DRUG RESISTANCE. MEDIATES THE TRANSLOCATION
CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
CC HEPATOCYTE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DISEASE: DEFECTS IN ABCB4 ARE A CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS TYPE IIT (PFIC), A FORM OF AUTOSOMAL
CC RECESSIVE LIVER DISORDERS, CHARACTERIZED BY EARLY ONSET OF
CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
CC ADULTHOOD.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M23234; AAA36207.1; -;
DR EMBL: X06181; CAA29547.1; -;
DR PIR: JS0051; DVHJ3.
DR PIR: AA2213; A42213.
DR HSSP: P13569; INBD.
DR Genem: HGNC:45; ABCB4.
DR MIM: 171060; -;
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABCtranspfltm.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR Prodom: PD00006; ABC_transpote; 2.
DR SMART: SMO0382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multimer family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 123 78 POTENTIAL.
FT TRANSMEM 153 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 354 POTENTIAL.

FT DOMAIN 355 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.
FT TRANSMEM 854 873 POTENTIAL.
FT TRANSMEM 937 956 POTENTIAL.
FT TRANSMEM 976 993 POTENTIAL.
FT DOMAIN 994 1279 POTENTIAL.
FT NE_BIND 429 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NP_BIND 1069 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT REPEAT 1 640 V -> VEVDFGFR (IN REF. 2).
FT REPEAT 641 1279 ATP (BY SIMILARITY).
FT CARBOHYD 91 91
FT CARBOHYD 97 97
FT CONFLICT 1093 1093
FT SEQUENCE 1279 AA; 140682 MW; 3D58C98B5C8B6087 CRC64;

Query Match 35 48; Score 2379.5; DB 1; Length 1279;
Best Local Similarity 40.88; Pred. No. 1,3e-132;
Matches 340; Conservative 231; Mismatches 466; Indels 87; Gaps 17;

QY 42 DLAPSTTATPADKEHREPKSSNNNAVSEVDALIAHLPEDROVLTGLEETVNIS 101
DB 2 DLBAKNGRA-----WRPTSAEGDEFGISS-----KORRKYT---VAM-IG 41
QY 102 FGLMRYATKMDILIVISTICAIAASTFORIML-YQISYDFEYD----- 146
DB 42 VLTFLFRSPQODKLFMSIGITAIHAGSGPLAMIVFGEKTDKFDVTAQNSFPVNFSS 101
QY 147 -----ELKNLYFVYLGIGEFVYVYSTVGFITYTGEATOKIRBYLYESTLIRON 196
DB 102 LNPGLILEEMRYVAYVYSGAGVLAAYIOVSFWTLAAGROIKRQKFEHALROE 161
QY 197 IGYFDKLGAGEVYTRITADTNLIDGISEKVGTLTALATFVAFIAYVKKWKLILCS 256
DB 162 IGFWDINDTTELMTLRITLDISIKSEIGDKVGFPAVATFFAGFVIRGKMLLVIM 221
QY 257 STVALVLTMGSGSOFIITYSKSLDSYGAGTVAEYVSSIRNATAFSTODKLAKOYEV 316
DB 222 AISPILGLSAAVWAKILSAFSDKEILAAVAKAGVABEALGAIPTVAFGGONLEERYOK 281
QY 317 HIDEAEKWKGNQIVYFNGAMFGMYSNGYGFNGSRFLVDGAVDGDILTVLMTL 376
DB 282 HLENAKEITIKRAISNIMWGIAFLDLYASVLAFAVGGTLYISKYTTIGMNTVFSTL 341
QY 377 ISFSISLGNVSPPNAQFTNVAANAIFGTIDROSPLDPSNBEKTLDFHEGHJELRNVAH 436
DB 342 IGAFSVGAAPCIDAFANRGAAYVFDLIDNPKIDSPSEGHKPDSTIKGNEFDVNF 401
QY 437 IYSPREYIVMEDVSLSPAGKTTALVSPSGSKSTVGLVERFYWPVPGTVLLDGHDIK 496
DB 402 STYSRANVRIKLGMLKQVSGQTVLVGSSGCKSTTVQLIQLRXPDEGTINIDGDIR 461
QY 497 DLNLRLROOISLVSOEPLVFGTTIYKINRHLIGIKYNEBEDKVARELLENAAAMAH 556
DB 462 NRVNVTLRITIGVSOEPLVSTTIAENICYG-----KGVNVTDEIK---AVEANAY 512
QY 557 DFTALPEGETYVNGSGRGLVSGGOKORIAIARAVYSDPKILLDEATSAIDTKSEGVYQ 616
DB 513 EFTMLKPOKFDVLVERGQVSGGOKORIAIARALVYRNPKILLDEATSAIDTKSEAEVQ 572
QY 617 AALERAAGRTTIVIAHRLSTIKTANIVLVNGLIAGQTHDELVDNRGAYRKLVEAQR 676
DB 573 AALDRAREGRTTIVIAHRLSTVRNADVLAGFEDVYIEGSGHSELMKKEGVYFKLVNQT 632
QY 677 INEQKEADALEADADLTNADIAIKTASASSSLDQKPTTIDRTGHSVSSAIIISKR 736
DB 633 SGSOIQSEFE-----LNDEKATRMADP-----NWKSRSLFRHSSTOKNLKNSQMCQK 679
QY 737 PPE-----TTPKYSIMTLTKFVASFNRPETIPMLIGVSVLAAGGQPTQAVLYAKA 788
DB 680 SLIDVETDGLVANVPVSLKYLK-----LNKTEWPFYVVGVCALANGGLQPAFVSIFSEI 735

```

QY 789 ISTLSPEOSYKRLHDAFWSLMFFVVGIIOTITSTNGAFAVCSERLIRARSTAFR 848
DB 736 IAFG-PGDVAWK-OOKKNFISFLFELGISEFTEFLQFTGKAGEILTRRLRSMWAK 793
QY 849 TILROIAFEDKRENSGTALTSFSLSTETKHLSSVSGTTLGTLMTSTTGAAITIALAIG 908
DB 794 AMLRQMSWMDHKNSTGALSTRLATDAVQCATGTRLLIAQNTIANLTGTIISFIYG 853
QY 909 WKALVCISVVPVLLACGFREFYMLAOFOSRSKLAVEGSANFACATSSIRTVASLRLER 968
DB 854 WQTLTLLLANVPIIVASGIYEMKLAGNMKRDKELEAKKITEALENTRTVYSLTQER 913
QY 969 DWEIYHAOLDAGRTSLISVLRSSLLYASSQALVEFCVAGFWYGGTLL--GHHEX-DI 1025
DB 914 KESMTVEKLYGPRNSVQKANHIGITFISISQAFMFYSYAGCFRFGALIVNGHMRRDY 973
QY 1026 FRFVCFSELTFFAOSAGVYFSFAPDMGRKNAAEFRRLPFDKRPQIDNNSEGEKLEY 1085
DB 974 --TIVFSALVFAGVALGHAASSFAPDYAKAKLSAHLFMLEFROPLIDSYSEGLKPKDF 1030
QY 1086 EGEIERNVHFRPTPEQPLRGDLTVKPGQYVALVSGSGKSTTIALLEFYDIA 1145
DB 1031 EGMTNEVENFPTANPVYLOGLSLEVKKGTIALVSGSGKSTVVOLEFRIDPLA 1090
QY 1146 GSTLVGDKDISKLINSYRSFSLVSGEPTLYOGTIKENILLGIVEDVPEELIRAKXD 1205
DB 1091 GTVLDDGEAKRLNVOMLRQGLIVSOEPLIFDCSIENIAYGDNRSVQDEIVSAKA 1150
QY 1206 ANIYDIMSLEPFNTVYSGKGMLSGGQKQRAVIALRLDRKILLDQATSLDSESE 1265
DB 1151 ANIHPETELPHKRYETRVGKGQTLSSGQKQRIALIAALROQIILLDQATSLDSESE 1210
QY 1266 KVVQALDAAARGRTTIAVHARLSTIQKADVIYVFDQKIVSGTHSELVOKGRYELV 1325
DB 1211 KVVQALDAAARGRTTIAVHARLSTIQKADVIYVFDQKIVSGTHSELVOKGRYELV 1270
QY 1326 NLOS 1329
DB 1271 SVQA 1274

RESULT 10
MDR2_RAT
ID MDR2_RAT STANDARD; PRT; 1278 AA.
AC 008201;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 2 (P-glycoprotein 2).
GN ABCB4 OR PGY2 OR MDR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer;
RX MEDLINE=93376516; PubMed=8103593;
RA Brown P.C., Thorgerirsson S.S., Silverman J.A.;
RT Cloning and regulation of the rat mdr2 gene.;
RL Nucleic Acids Res. 21:3885-3891(1993).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: L15079; AAA02937.1; -
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABC_transport.
DR Pfam: PF00664; ABC_membrane_2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 709 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 710 730 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 831 851 POTENTIAL.
FT TRANSMEM 852 872 POTENTIAL.
FT TRANSMEM 935 955 POTENTIAL.
FT TRANSMEM 972 992 POTENTIAL.
FT DOMAIN 993 1278 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (POTENTIAL).
FT NP_BIND 1068 1075 ATP (POTENTIAL).
SQ SEQUENCE 1278 AA; 140655 MW; DAFB6BE745AF73BF CRC64;

```

```

Query Match 35.1%; Score 2357; DB 1; Length 1278;
Best Local Similarity 40.9%; Pred. No. 2.8e-131;
Matches 524; Conservative 224; Mismatches 464; Indels 68; Gaps 14;

QY 89 LKQLEIRKIVNISPFQ--LMRYATKMDILIMVSTICATAASTFORIML-YOISYDF 144
DB 23 ISNQSREKKKKVNLIGPLTLEFRYSMDQDKLFMLGTMALAHSGELPMHIVEEMDKF 82
QY 145 YD-----ELTKNLYFYVLGIGEFYTVVSVGFITYGEHA 180
DB 83 VDNAGNFSLPVNSLSMLNPGRILEEEMTRAYVYSGIGGVLLAAAYIOVSFWTLAAGRQ 142
QY 181 TOKIREYLLSILRONIGYEDKLGAGEVTRITPDTNLIDGISEKYGTLTLATFVTA 240
DB 143 IRTKROFFHAILROEKGMFDIKGTLENLRLDDISKISEGIDKGMFFQATATFPAG 202
QY 241 FIATVYVWKIALCSSTIALVLTMGSGOFTIKYSKSLDSYAGAGTVAEEVYSIRN 300
DB 203 FIYGFIRGKLLVITVMTAILGLSTAVWKKILSTFSDKELAAVAKAGVAEELGAIKT 262
QY 301 ATAFGTODKLAKOYEVLDEAEKMGTRNQIYMGFMIGAMGLMYSNTGLGFMGSRFLVD 360
DB 263 VIAFGGNKELERYOKHLENAKKIGIKKALSANISMGIAFLIYASVALFWYGSTLVIS 322
QY 361 GAVDVGLITVIAILLGSGSIVGNVSPNAQFTTAAVAAAKITGDTIDRQSLDPSNEK 420
DB 323 KETIIGAMTVFSTILGASVSGAAPCIDFPNARQAAYVIFIDINNPXIDFSERGH 382
QY 421 TLDFEGLHILRNKHTYPSRPEVTVMEDVLSMPACKTALVGPSSGKSTVGLVERF 480
DB 383 KPDSIKKNLESDVHFYPSRANKIKLGNLKKYSQYVALVNSGCGKSTVQQLRL 442
QY 481 YMPYGVLLDGDHIDKLNLEMLRQOISLVSOEVLFGTTIYKNIRGLIGTKYENSED 540
DB 443 YDPEGTISIDGDIRNFNKNCLFEFIVGVSOEVLFTTIAENIRG-----RGNTYMD 497
QY 541 KYRELINAKMANAHFITALPGEYTNVGOBFLISGQKQRIATARAAYVSPKILL 600
DB 498 EIKR---AVKENAAYDFIKLTPKFDLVGDRQAQSGGQKQRIATARAAYVSPKILL 553
QY 601 DEATSLDFTKSEGVQAALERAAGRTTIVIAHRLSTIKTAHNIIVLVNCKIAEGTIDE 660

```

```

DB 554 DEFTSALDTSSEAVQALDARERTIYIAHRLSTVRADYAGPEDGYIEQSHSE 613
OY 661 LVDRGAKRLVAPORINEQKADADADADLTNNADIKITYASSASSDDKPTTID 720
DB 614 LIKEGTYFLVMQTSQSIILSEEEF-----VELSDERAKAGVAD--NGMKARIF 662
OY 721 RTGTHKSVSA-----ILSKRPETPKYSILMTLLKRVASPNREIYMLIGLVSV 772
DB 663 RNSKTSKSSRAHONRLDETVELDANVPVPSFLKLR---LNKTEWYFVVGTCAT 718
OY 773 LAGGQPTQAVLYAKAISTLSPESQSKLRHDAFWSLMEFVVGIIQFTQSTNGAFA 832
DB 719 ANGALPAPFIIISEMIAIRG-PGDDTVK-QQCNMFSLVFLGHSFFTLQGTGTF 776
OY 833 VCEBRLIRARSTAFRTIIRODIAFPDKENSGLALSTLSTETKHLGVSGLTILM 892
DB 777 KAGEIILTRLSMAFKAMFQDMQSWDFDHKNSGALSTRLATDAQVAGATGRLALIA 836
OY 893 TSTLGAIIIALIGMKIALVCISSVYPVLLAGFYRFYLAQFQSKSLAEGSANFAC 952
DB 837 NTAINDGTIISTFYGMQLLLSVYPTAVAGIYEMKMLAGNAKDKKEMAGRIAT 896
OY 953 EATSIRTVASLTRERDWEIYHAOLDAQRTSLISYRSLLYASSQALVFCVALGFY 1012
DB 897 EAIENTRTVSLTQERKESEMYEKLHGPYRNSVRKAHIYIGFISQAFMYTSYACFR 956
OY 1013 YGCTLL--GHHEV-DIFRFVCSSEILFGAQSAGTVSEAPDNGKAKNAAEFRRLPDR 1069
DB 957 FGSYLLVNGHMRKDV---ILVFSAYLVGVALGHASSFPDYAKAKLSAAYLFSEERQ 1013
OY 1070 POLDNMSSEGEKLEVEGEIEFRNVRYPTRPEQVLRGLDITVPGQVVALVSGCG 1129
DB 1014 PLIDTSRGEMADKEGSTEVEVNFPTANVPVGLSLLEVKKQGLVALVSSGCG 1073
OY 1130 KSTTIALREFYDAGSILVYDKDISKINSRSLVSQEPLOYGTIKENILGI 1189
DB 1074 KSTVQLLREFYDPMAGTVLLDQGAOKLVQMLRGVIGSEPIIFDQSIKNTAYGD 1133
OY 1190 VEDDVPEELIRAKRANITDFIMSLPEGFNYVSGKGLSGGQORVAIARALIDPK 1249
DB 1134 NSRWVSQDIDIVRAKANHPEITLPKRYETRVGDKGTQLSGGQORAIARALIDPK 1193
OY 1250 ILLDPAATSESEKVVQVQALDAAARGTIAVAHRLSTOKAVIYVFDGKIVESG 1309
DB 1194 VALLDPAATSESEKVVQVQALDAAARGTIAVAHRLSTIONADLIIVDNGKVERG 1253
OY 1310 THSELVQKRGYYELVNIQS 1329
DB 1254 THQOLLAGKGYFSMVNIQA 1273

RESULT 11
MDR3_CRIGR STANDARD: PRT: 1281 AA.
AC P23174:
DF 01-NOV-1991 (Rel. 20. Created)
DF 01-NOV-1991 (Rel. 20. Last sequence update)
DF 16-OCT-2001 (Rel. 40. Last annotation update)
DE Multidrug resistance protein 3 (P-glycoprotein 3).
GN PGY3 OR PGP3.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
RL DNA Seq. 2:89-101(1991).

```

```

CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M60042; AAA68865.1; -.
CC HSSP: P13569; INBD.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR003439; ABC_transport.
CC InterPro: IPR001140; ABCtransportTM.
CC Pfam: PF00005; ABC_tran; 2.
CC Pfam: PF00664; ABC_membrane; 2.
CC ProDom: PD000006; ABC_transport; 2.
CC SMART: SM00382; AAA; 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
CC Multigene family.
CC DOMAIN 1 57
CC FT TRANSMEM 58 78 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 122 142 POTENTIAL.
CC FT TRANSMEM 191 211 POTENTIAL.
CC FT TRANSMEM 218 238 POTENTIAL.
CC FT TRANSMEM 299 319 POTENTIAL.
CC FT TRANSMEM 328 348 POTENTIAL.
CC FT DOMAIN 349 712 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 713 733 POTENTIAL.
CC FT TRANSMEM 758 778 POTENTIAL.
CC FT TRANSMEM 834 854 POTENTIAL.
CC FT TRANSMEM 855 875 POTENTIAL.
CC FT TRANSMEM 938 958 POTENTIAL.
CC FT TRANSMEM 975 995 POTENTIAL.
CC FT DOMAIN 996 1281 CYTOPLASMIC (POTENTIAL).
CC FT NP_BIND 429 436 ATP (POTENTIAL).
CC FT NP_BIND 1071 1078 ATP (POTENTIAL).
CC SEQ ID 1281 AA; 140866 MW; 2203EF61EBB29602 CRC64;

Query Match 35.1%; Score 2357; DB 1; Length 1281;
Best Local Similarity 41.3%; Pred. No. 2.9e-131;
Matches 526; Conservative 224; Mismatches 449; Indels 76; Gaps 13;

OY 97 KVN-ISFGLMRYATKMDILIMVITICAAASFORIML-YQISYDEFY----- 145
DB 36 KYNLIGPLTLFRSQMDQLFMILGTIMAIHSGSLPLMIYFGEHTDFVNNAGFSLP 95
OY 146 -----DELTKNLYEYVYLGIGEFVYVSTVGFIYTGHAQKIREYLE 190
DB 96 VNFSLIMINPGRILREEMKRYAVYSGLGGLVYAVYQVFWTLAAGQIKKIRQNFH 155
OY 191 SILRQIGFEDLAGEVYTRITADPNLQDQISEVGLTIRALATFTVAFTIAYKYTK 250
DB 156 ALTRQDMGFMDIKGTETLTRLDDISKISEGIGDVGAFQAVAFVFGFTVGFRGK 215
OY 251 LALICSTIYVALVLTWGGSGQFLIKSKSLDSYGAGTVAEVISSINATAFGQDGL 310
DB 216 LTLVYMAISPIILGSAVAWAKILSTFSDKELAAYAKAGNAVEALGATVYAFGQNK 275
OY 311 AKOYEVHLDAAEKWGKTKNOIVGFMIGANFGLMTSYNGLGFMGSRFLVDGAVDGLT 370
DB 276 LERYQKHLENARKIGIKKAISANISWGIAPFLIYASYALAFVYGSPLVSKETYGNAWT 335
OY 371 VMALITGSFSGNVSPPNQAFTNVAAMAAKTFGTIDRQSPDIPYSNEKTLDFHEGHE 430

```

```

Db 336 VEPISILIGAFSGVQAAPCIDAFANARGAAYIEDIDNNPKIDFSERGHKPKDSIKGNLD 395
QY 431 LRRNKHLYPSRPETVMEVSLSPACKTATVSPSGSGKSTVGLVERFPMYRGTVLL 490
Db 396 FSDVHSYSPRANIKIKLIGKLVQSGQYVALVNSCGGTTTLQILQRLYDPEGIISI 455
QY 491 DGHDIKDLNRLMLRQOISLVQSEVLEFGTTTYKNIRHGLIGTYENESDKVRELIENAA 550
Db 456 DGODIRNFNRVRLREITIGVVSQSEVLESTTIAENIRFG-----RGNVMEIKKAV 506
QY 551 KMANNDFTALBEGYTNGVSGFLLSGGOKRIALARA VSPKLLILDEATSALDTK 610
Db 507 KEANAEEFIKMLPQKEFTLVGERGAQLSGGOKRIALARA VSPKLLILDEATSALDTK 566
QY 611 SEGVQALERRAAGRTTYIARHLSITKTAHNVLVVYNCKIAKAOGHDELVDGAYRK 670
Db 567 SEAEVQAALDKAREGRTTYIARHLSITKTAHNVLVVYNCKIAKAOGHDELVDGAYRK 626
QY 671 LVE-----AQRINEQKADALEDAEDLTNADIATKRTASASSSDLDGKPTTIDRTGTH 725
Db 627 LVNMQTSGSOLISQEEFEVLESEKADGMTP-----NGMKSHIFRNSTK 670
QY 726 KVSVA-----ILSKRPETTPKSLMTLLKFAVSPNRPETPYMLIGVSVLAGGC 777
Db 671 KSLKSSAAHHHRLVDADDELIDANVPVSLKVLK---LKTETPVFVGVCAIVNGAL 726
QY 778 QPTQAVYAKAISTLSPEQSOKLRHDADFWSLMFVVGIIOTITOSTNGAFAVCSER 837
Db 727 QPASTILSMAIFG-PGDDANK-QOKMLFSLVFGIGLVSTFTFLQGTGKAGEI 784
QY 838 LIRRASTAFRTLLRODIAEFDEENSTGALSTSLSTETKHLSCVSGVLTGTLMTSTLL 897
Db 785 LTRLRSMARKAMLRQMSWFDYKNSTGALSTRLADRAOVQATGTRILALINQNTANL 844
QY 898 GAAITIALAGMKIALYCVSVPVLLACGFYRFLMLAOPSRKSLAEVSGSNPCETSS 957
Db 845 GTGIIISIFYQWOLTTLLSVPPVIAVSGIVEMKMLAGNKRDKKALEAGKTAIEN 904
QY 958 IRTVASLTERDWEIYHQAOLDAGRTSLISVLSLLYASSQALVFCVAGLFWYGTLL 1017
Db 905 IRTVASLTERDWEIYHQAOLDAGRTSLISVLSLLYASSQALVFCVAGLFWYGTLL 964
QY 1018 L--GHNEY-DIFREYVCFSEILFGAQSAGVTFSPADPMGAKNAAEFRLLDRKPOIDN 1074
Db 965 IYVGHMFRRDY---ILVFSALVFGAVALGHASSFAPDYAKAKLSAHLFSLERQPLDS 1021
QY 1075 WSEGEKLEVEGEIEIRANHERFRPRRBPVLRGLDLYKPGQYVALVYPSGGKSTTI 1134
Db 1022 YSGGLMPDPFEGSVTFNEVEFNPPTANMPVLQGLSLEVKKGOTLALVSSGGKSTVV 1081
QY 1135 ALLERFYDAISILVDGKDISKLINSYRSFLSVSQEPFLYOGTITKENILGIVEDV 1194
Db 1082 QLLERFDPAGVYLLDQGAKKNTIOMLRAQGLVSOEVLFPCCSLAENIAYDNRV 1141
QY 1195 PEEFLKACKDANIYDIMSLEPEGNTVSGSGKMSGGOKORAVARALALDRPKILLD 1254
Db 1142 SOEIVVAAAANHPIETLPQKYKRVGDKFTQSGGOKORALIRRALIRQRRVLLD 1201
QY 1255 EATSALDSESEKVVQAALDAARGRTTIANHRLSTQKADVYVYPOGKIVESGISEL 1314
Db 1202 EATSALDSESEKVVQAALDAARGRTTIANHRLSTQKADVYVYPOGKIVESGISEL 1261
QY 1315 VOKGRYELVNLOS 1329
Db 1262 LAQKGIYFSMNTQA 1276

```

```

RESULT 12
AB11_HUMAN STANDARD; PRT; 1321 AA.
AC 095342: 09JUN82:
DF 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11).
GN ABCB11 OR BSEP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS PFC2 GLY-297; GLU-461; GLY-482;
RP ARG-982; CYS-1153 AND GLN-1268.
RX MEDLINE=99021377; PubMed=9806540;
RA Strautnieks S.S., Bull L.N., Knisely A.S., Kocoshis S.A., Dahl N.,
RA Ansell H., Sokal E., Dahan K., Childs S., Ling V., Tanner M.S.,
RA Regalwalla A.F., Nemeth A., Pawlowska J., Baker A., Miel-Vergani G.,
RA Feimer N.B., Gardner R.M., Thompson R.J.;
RT "A gene encoding a liver-specific ABC transporter is mutated in
RT progressive familial intrahepatic cholestasis."
RL Nat. Genet. 20:233-238(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Mol O., Hooiveld G.J.E.J., Jansen P.L.M., Muller M.;
RT "Cellular localization and functional characterization of the human
RT bile salt export pump (BSEP).";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT ALA-444.
RX MEDLINE=21686803; PubMed=11829140;
RA Saito S., Iida A., Sekine A., Miura Y., Ogawa C., Kawauchi S.,
RA Higuchi S., Nakamura Y.;
RT "Three hundred twenty-six genetic variations in genes encoding nine
RT members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the
RT Japanese population."
RL J. Hum. Genet. 47:38-50(2002).
RN [4]
RP VARIANTS INTRAHEPATIC CHOLESTASIS LEU-284 AND ASP-1004.
RX MEDLINE=21674594; PubMed=11815775;
RA Chen H.-L., Chang P.-S., Hsu H.-C., Ni Y.-H., Hsu H.-Y., Lee J.-H.,
RA Jiang Y.-M., Shau W.-Y., Chang M.-H.;
RT "PFC1 and BSEP defects in Taiwanese patients with chronic intrahepatic
RT cholestasis with low gamma-glutamyltranspeptidase levels."
RL J. Pediatr. 140:119-124(2002).
CC - FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
CC INTO THE CANALICULUS OF HEPATOCYTES.
CC SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
CC SITU (BY SIMILARITY).
CC - DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC ATP BINDING CASSETTE (ABC) DOMAIN.
CC - DISEASE: DEFECTS IN ABCB11 ARE THE CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS 2 (PFC2), AN INHERITED LIVER DISEASE OF
CC CHILDHOOD. PFC2 IS CHARACTERIZED BY CHOLESTASIS AND NORMAL SERUM
CC GAMMA-GUTAMYLTRANSFERASE ACTIVITY. DEFECTS IN ABCB11 ARE ALSO
CC FOUND IN CASES OF CHRONIC INTRAHEPATIC CHOLESTASIS WITHOUT OBVIOUS
CC FAMILIAL HISTORY OF CHRONIC LIVER DISEASE.
CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF091582; AAC7455.1; -.
CC EMBL; AF136523; AAD28285.1; -.
CC Genew; HGNC:42; ABCB11.
CC MIM; 603201; -.
CC MIM; 601847; -.
CC InterPro; IPR003593; AAA_ATPase.

```

DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABC_transport.
DR Pfam: PF00005; ABC_tran. 2.
DR Pfam: PF00664; ABC_membrane. 2.
DR ProDom: PD000006; ABC_transport. 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transmembrane; Transport; Disease mutation; Polymorphism.
FT DOMAIN 1 62
FT TRANSMEM 63 83
FT TRANSMEM 84 147
FT TRANSMEM 148 168
FT TRANSMEM 169 215
FT TRANSMEM 216 236
FT TRANSMEM 237 240
FT TRANSMEM 241 261
FT TRANSMEM 262 319
FT TRANSMEM 320 340
FT TRANSMEM 341 353
FT TRANSMEM 354 374
FT TRANSMEM 375 755
FT TRANSMEM 756 776
FT TRANSMEM 777 794
FT TRANSMEM 795 815
FT TRANSMEM 816 869
FT TRANSMEM 870 890
FT TRANSMEM 891 911
FT TRANSMEM 912 979
FT TRANSMEM 980 1000
FT TRANSMEM 1001 1011
FT TRANSMEM 1012 1032
FT TRANSMEM 1033 1321
FT NP_BIND 455 462
FT NP_BIND 1113 1120
FT CARBOHYD 109 109
FT CARBOHYD 116 116
FT CARBOHYD 122 122
FT CARBOHYD 125 125
FT CARBOHYD 284 284
FT VARIANT 297 297
FT VARIANT 444 444
FT VARIANT 461 461
FT VARIANT 482 482
FT VARIANT 982 982
FT VARIANT 1004 1004
FT VARIANT 1153 1153
FT VARIANT 1268 1268
FT CONFLICT 339 339
FT SEQUENCE 1321 AA; 146392 MW; D44ACACC48DEA371 CRC64;
Query Match 34.6%; Score 2328; DB 1; Length 1321;
Best Local Similarity 38.7%; Pred. No. 1.5e-129;
Matches 508; Conservative 237; Mismatches 469; Indels 100; Gaps 14;
QY 90 KTQLEIRK-----VNISFGLMRYATKMDLILINWISTICAAASTFORIML----- 136
DB 31 KSRLODEKGDGVRGFFQLFRSSSTDLIMFVGSICAFHIGIAOPVILLIGTMTDPF 90
QY 137 --YQSYDEFY-----DELTKNLYEYVLGI 160
DB 91 IDYDELQELQIPGKACVANNITWNTNSLQNMNTNGTRCGLINIESEMIRKFSYAGIAV 150

QY 161 GEFYTVYVSTVGFIYGEHATQKIREYYLESILRONIGYDCKAGVYTRITADNLQ 220
DB 151 AVLTGTIGTCFVYIAAARQIQKRRKFRIMMELGMDCSVSGELNTRFSDIDKIN 210
QY 221 DGISEKVGTLTALATFVAFTIAYVYKMLALICSSITVALVYTMGGGSOFT----IKY 276
DB 211 DALADQALALIQNTMTISGIFLIGFFGKMLTY-----IISVPLIGTATIGLSVKF 266
QY 277 SKSLDSYGAGGYAAEVEISSIRNATAFGTQDKLAKOYEHLDEAEKMGTRKNOYGMFI 336
DB 267 TDYELKAYARAGVAADEVISMRTVAFAFGKEKEVEYKENVFAQMGIRKGIYMGFT 326
QY 337 GAFPGIMYSYKCGFWMGSRFLVD--GAVDVGDIITVIMALLISFSLGNVSPNAQAFNA 395
DB 327 GFVWCLFELCYAAVAFWGSFLVLDGEGYTGTVQVLEISYVAGALNIGNASPCIEAFATG 386
QY 396 VAAAKIFGTIDRQPLDPYNSNEKTLTDHDEGHLELRNHYIPSRREVYVMDVSLMP 455
DB 387 RAATSTIFETIDRKPIIDCSBEGYKIDRIKGELEIRNVTFFHPSRPEVAILDNNVYK 446
QY 456 AGRTTALVPGSGSGKSTVGLVERFYMPVRYGTVLLDGHDKLNLMLRQQLSVSQEPY 515
DB 447 PGEMTALVPGSGAKSTALQIQFYDPCGAVTVGDHDIRSLNIQMLRQIGIVQEPY 506
QY 516 LFGTTIYKINRBLIGIKYENESDKVRELTEMAAKAANADHTTALPGSETNQGQGF 575
DB 507 LFSITTAENIRYG-----REDATMEDIVQAANKANNYNFIMDLPOQFDPLVGEGG 557
QY 576 LLSGQKORAIARAVYSDPKILLDEATSAULTKSGVQAALERAABERTIYVAHRL 635
DB 558 QMGCGQKORAIARALIRNPKIILLDMATSALENSBAMQEVLSKIQHHTIISVAHRL 617
QY 636 STIKTAHNIYVAVNGKIAEGTDELVDRCGAYRKILEADQINEOREADALDEADLT 695
DB 618 STYRADTIIFGFGHGAVERGTHEELLERKGYFTLVITQSQNQ---ALNEDIKDAF 673
QY 696 NADI-----AKITAS--SASDLDKRP--TTIDRTGTHS--VSSAILSK 735
DB 674 EDMILARTSRGYSQDSLRSIRQSRKSQLSYLVHPEPLAVVDHKSRYEDRDKDIPVO 733
QY 736 RPETTPKYSIMLTLKFVASFNRPEIPLYMLIGLVFSLVAGGGQPTQAVIYAKAISTSLP 795
DB 734 EHEEPAP-----VARILKFSAPMPMYLVGSAVANGTVPLPAFLFSQLGTFPSIP 786
QY 796 ESQYSLRHADAFWSLMFVVGIIQPTOSTNGAARVCSERLIRARSTAFETIIRODI 855
DB 787 DKE--RORSQINGVCLLFVAMGCVSLFTQFLQYARAKSGELLTKLRKFRFAMLGDOI 844
QY 856 AFEDKRENSGATSLSTETKHLKLSGVSQVTLGTTIMSTLTGALITIALAIGWKALVC 915
DB 845 AMFDLIRNSPGLATTLRLADASQVGAAGSQIGIMVNSFTNVTYVAMIIAFSFSWKLSTVI 904
QY 916 ISVVPVLLACGFYRFYMLAFOGSRKLAYESGANFACATSSIRYVASILREDDWEIYH 975
DB 905 LCFEPLLASGATQTMILGFPASRDQALEMVGQITNEALSIRTYAGIGKERRFTEALE 964
QY 976 AQLDAQRTSLISLVKSSLLVYSSQALVFECVALGFWGTLGHHEHYDIFFEYCFSEI 1035
DB 965 TELEKEPFTAIQKANIYGCFAFAQICIMFTANSASRYGVIILNSGHLFSYFRTISAV 1024
QY 1036 LFGAQSAGTVFSPAPDMGAKAKNAAEFRRLFDKRPQIDMMSSEGEKLEFVEGIERNVH 1095
DB 1025 VLSATVLRGAFSTPYPAKAKISAAFFQILLDRPISYVNTAGEKMDNFQKIDFVDC 1084
QY 1096 FRPPTRPQPVLRGDLTVFKPGQYVAVLVPGSGGKSTTALLERFDALAGSLVDGKXI 1155
DB 1085 FTYPSRPDQVNLGSLVSLSPGOTTLAFVSSGCGKSTSLQLERFYDPOGVMYDGHOS 1144
QY 1156 SKLINYSRFLSLVSOEPTLYOGTITKENITLIGIYEDVDPEEFLLKACDANIYFIMSL 1215
DB 1145 KKVNVQFLNSNIGYQSEVPLFACSIMDKIKGDNKKEIPMERVIAAQAQDLHPVMSL 1204
QY 1216 PEGFNTVYSGSKGMLSGGQKORVATIRALLRDPKILLDEATSAIDSESEKVVQAALDA 1275

DB 1205 PEKYTEINWSSQSGSOLSRGEQRIALAAVYRDKILLDQATSDALDSEKTYQVALDKA 1264

OY 1276 ARGRTIAVAHRLSTIOKADVIYVPOGKIVEGTSSELVOKGRVYELNLOS 1329

DB 1265 RESRTOIVAHRLSTIONADIIIVMAGVIEGTHHELMAGKAYKLVTTGS 1318

RESULT 13

AB11_MOUSE STANDARD: PRT; 1321 AA.

AC 090Y30: 090ZEB: 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)

GN ABCB11 OR BSEP OR SPCP.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=1090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=20076398; PubMed=10607905;

RA Green R.M., Hoda F., Ward K.L.;

RT "Molecular cloning and characterization of the murine bile salt export pump";

RL Gene 241:117-123(2000).

RN (2)

RP SEQUENCE OF 463-635 FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Liver;

RA Salkar R., Suchy F.J., Ananthanarayanan M.;

RT "Molecular cloning of mouse liver bile salt export pump (bsep).";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBD databases.

CC - FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS INTO THE CANALICULUS OF HEPATOCYTES.

CC - SUBCELLULAR LOCATION: Integral membrane protein.

CC - TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN SITU.

CC - DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES, EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN ATP BINDING CASSETTE (ABC) DOMAIN.

CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb.slb.ch).

CC -----

DR EMBL: AF133903; AAF14372.1; -

DR EMBL: AF186585; AAD56419.1; -

DR MGD: MGI:1351619; Abcd11.

DR InterPro: IPR003439; AAA_ATPase.

DR InterPro: IPR003439; ABC_transporter.

DR InterPro: IPR001140; ABCtransprtTM.

DR Pfam: PF00005; ABC_tran; 2.

DR Pfam: PF00664; ABC_membrane; 2.

DR ProDom: PD000006; ABC_transportr; 2.

DR SMART: SM00382; AAA; 2.

DR ProSite: PS00211; ABC_TRANSPORTER; 1.

KW ATP-binding; Transmembrane; Transport.

FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 63 83 POTENTIAL.

FT DOMAIN 84 147 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 148 168 POTENTIAL.

FT DOMAIN 169 215 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 216 236 POTENTIAL.

FT DOMAIN 237 240 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 241 261 POTENTIAL.

FT DOMAIN 262 319 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 320 340 POTENTIAL.

FT DOMAIN 341 353 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 354 374 POTENTIAL.

FT TRANSMEM 375 755 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 756 776 POTENTIAL.

FT TRANSMEM 777 794 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 795 815 POTENTIAL.

FT TRANSMEM 816 869 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 870 890 POTENTIAL.

FT TRANSMEM 891 911 POTENTIAL.

FT TRANSMEM 912 979 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 980 1000 POTENTIAL.

FT TRANSMEM 1001 1011 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1012 1032 POTENTIAL.

FT TRANSMEM 1033 1321 CYTOPLASMIC (POTENTIAL).

FT NP_BIND 455 462 ATP (POTENTIAL).

FT NP_BIND 1113 1120 ATP (POTENTIAL).

FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 481 481 L -> P (IN REF. 2).

FT CONFLICT 633 633 T -> V (IN REF. 2).

SO SEQUENCE 1321 AA; 146675 MW; 15B5EBF15D32967 CMC64;

Query Match 33.4%; Score 2247; DB 1; Length 1321;

Best Local Similarity 38.1%; Pred. No. 9,4e-125;

Matches 502; Conservativity 246; Mismatches 456; Indels 112; Gaps 19;

OY 90 KTOLEEK-----VNISFGLMRVATKMDILIMVSTICATA--AASTQRMIVQISDER 144

DB 31 KSRLODKKKKGARVGFELFRSSKDNLMFMGVCALLHGMAGQGMIVFGILDF 90

OY 145 --YD-----ELKRVLYF--VYLGIGE 162

DB 91 VERDIERQELSPKVCNMNTIWINSPFNMNTNGSCGLVDINSEVIFSGIYAGV 150

OY 163 FTVV--YVTVGFIYTGHAATQKIREYLLSILRONITGYFDKLGAEVTRITADNLIQ 220

DB 151 AVLLGFGFOIRLWITGABOIRKKRKYFRIRIMEIGFDCSVGLNRFSPDINKID 210

OY 221 DGISEKGLTLTATVTAFTIAYVYKMLICSTIYALVITMGGSQFT----IKY 276

DB 211 EALADQALFLQRLSTLGLLGFYFGWKLTIV---ILAVSPLIGIAVIGLSYAKF 266

OY 277 SKSLDSYGAAGVAEEVSSIRNATAGTQDLAKQYEHVLDAAEKGRKNOITMGFMI 336

DB 267 TELELRKAYAKAGSTADDEVLSIRYVAFGENKEVEREKRLMAQMGIMKGMVGFPT 326

OY 337 GAMFGLMYSNYGLGFWMGSRFLVD--GAVDVGDILITVMAILIGSFGNVSPPNAQFTNA 395

DB 327 GYMWCLIFCYALAFWYGSRLVLDGEYTPGTLIQIFLCVIAAMNIGMNSCSEIFSTG 386

OY 396 VAAAKRFGITIDQSPIDPNSNEKTDHREGHLELRNKHITRSREVTYMEVSI.SMP 455

DB 387 CSAASSIFQITIDROPVADCKSGDGYKLDRIKGELEFINVTFHPYSRPEVILNLSVIX 446

OY 456 AGKTTALVPGSGSGKSVVGVVERFYMPVGVYVLDHDKDLRLRLOOISVSGEPV 515

DB 447 PGETTAVVSSGSGAKSTALDIIQAFYPCRCMWTLLDHDRLSRLRLOOIGIVDEP 506

OY 516 LFGTTIKNIRHGLIGTYENESEDKVRELLENAAKXANAHDTALPEGETVWVGQRF 575

DB 507 LFSITTIENIRLG-----REATVMDIYQAKRANVNFIMALPQGFDTLVGSGG 557

OY 576 LLSGGQRIARAAVSDRKILLDQATSDLTSGCVQALERAEGRTTIVIAHRL 635

DB 558 QMSGQKORVAIARALIRKPKILLDQATSDALDSESAKVGALNKRIOHGTITISVAHRL 617

Db 148 IGAIVLTGTYICFNGIAAHOIQMRKSYFRIMPMGIGWDCNSVGLINTPEFSDEN 207
 QY 218 LIODGISEKVLTLTALATEVTAIIAYVYKWLALICSSITVALVLTMGGSGOFI---- 273
 Db 208 KINDSADQALFIOGTSPILFGLVGEFSQMMKLTIV-----LISVPLIGALAIIGLSV 263
 QY 274 IRTSKSLDSYAGAGYIAEEVYISSIRNATAPGICODKLAKOYVHLDEAKWGKINQIYMG 333
 Db 264 SKETDELKAYAKAGSVADDEVISMRTVAAFGEKEVEREKVLVAORNGIKRGIVMG 323
 QY 334 FMIGAMGLVSNYGLGFMSGRF-LYDGAVDVGDILTVMALIIGSEFVNSPNOAF 392
 Db 324 FFGYIMCLIFCYALAFMGSKVLVEEGEYSPALVQIFLSTVIGALNMGASPCLEAF 383
 QY 393 TNAVAAAKIFGTIDROSPIIDPYNSNEKTLDEHGHIELRNKHIYSREPVYMEVSL 452
 Db 384 AAGRAAASSFEITIDRPIIDCMSEDOYKLERIKGEIEFNHVFHYHSREPVKILNLSM 443
 QY 453 SMPAGKTALVPSGSGKSTVGVIVFVYMPVRCVTLLDGHDKDLMRLRQOISLVSQ 512
 Db 444 VIRGEWALTALVPSGAGKSTALDILHRFYGTGEMVTVEHSHIDRSSHIOWLRNGIIVEQ 503
 QY 513 EPLVFTTIIKRNHIGLIGIKYENESDKYRELINAKAMAHDEFTALPEGYEVNQ 572
 Db 504 EPLVFTTIIKRNHIGLIGIKYENESDKYRELINAKAMAHDEFTALPEGYEVNQ 554
 QY 573 RGFLLSGGKORITAIARAVYSDPKILLDEATSAIDTSEGVQALERAEGRTTIVIA 632
 Db 555 GGGOMSGGKORAVAIARALRNKRIILLDMATSLDNESAMQALSKOHHGTIYVA 614
 QY 633 HRLSTIKTANHIVLVNGKTAEGCTHDELVDRCGAYRKLV--EAQR--INEQREADLE 687
 Db 615 HRPATITADVIIIGCEHGAVERGEELERKGVYFALVLOSRNQDGEENEKATE 674
 QY 688 D-----ADAEIDLNADIARIKTASSASSDLGKPTI-DRTGHKSSVSSAILSKRP 737
 Db 675 DDIPKFTSGNTOODSIRASLRQ-RSKQSLSYTAHEPPMAVEDKSTHEDRK---DKDL 730
 QY 738 P-ETTPKYSILWTLKLFVASFNRPEIYMLIGLVSYLAGGQPTQAVLYAKAISTSLP 795
 Db 731 PAGEDIEPASVVRIMK-----LNAPEMPYMLLGSKGAIVNGAVPLVFLFSQILGTSLP 786
 QY 796 ESQYSKIRHADRWMSLFVAVGIQITOSTNGAARVCSERRLRARASTAFRIILRODI 855
 Db 787 DKE--EORSQINGICLFLVTLGCVSFTFOGTYFAKSGELLKRLKRFEGFAMLCODI 844
 QY 856 AFEDKENSUGALTSFSTETKHLGSGVGTIITMTSTLGAIIIALAIGKIALVC 915
 Db 845 GWFDDLRNSFGALTRLATDASQVQATGSOIGMNVNSFTVVTAMIIAFLFSMKTGTGI 904
 QY 916 ISVVPVLLAGFYRYMMAFOFSKSLAYEGSANFACATSSIRTVASLIFRERDWEIYH 975
 Db 905 VCFEPFLALGALQTKMLTGFASRDQALEKAGIITSEALSNITVAGIGERKFIEFFE 964
 QY 976 AQDADGRTSLISVLRSSLLYASSQALVFCVALGFWYGGTLLGHHEYDIFRFVCESEI 1035
 Db 965 AELEKPKYKAIKANYGGLGFEFGSCITFTLANSASYRGYGLISNEGILHFSYFVRISAV 1024
 QY 1036 LFGNOSAGTVPSPADMGKAKMAAEERRLDPRPOIDNMBESEKLETVGELEFRVH 1095
 Db 1025 VLSATLALGRASSYTPYAKKAKISARFOLDROPINIVYSSAEKMDNFGKIDFVDC 1084
 QY 1096 FRYPTREPOVLRGLDITVFGQYVALVPGSGCKSTTIALLEFFYDAIAGSIIVDGDID 1155
 Db 1085 FTVYSRDIQVLANGLSVMSRQTLAVVSGSGCKSTSIQLEFFYDPDHQKWMIDGIDS 1144
 QY 1156 SKLINSYRSFSLVSEDEPTLYOGTICKENILGIVEDVPEEFLIKACKDANIYDFIWSL 1215
 Db 1145 RKNVIOQLRSNIGIYSEDPVIFACISIKDNIKYGDNTOEIPEERIIAAKKAQVHDFWSL 1204
 QY 1216 PEGNNTVYSGKMLSGGOKORAVAIARALLRDPKRIILLDEATSLDNESSEKVVQALDAA 1275
 Db 1205 PEKETVNTVSGSOLSHGEKORITAIARIVADPKRIILLDEATSLDNESSEKVVQALDAA 1264

QY 1276 ARGFTTAAHRLSTIQADYIVFDQKIVESGTHSELVQKGRYELVNIQS 1329
 Db 1265 REGRTCIYIAHRLSTIQNSDIIVAMSGOMVTEKGTHEELMAYQKAYKLVTTGS 1318
 RESULT 15
 ABIL_RAT
 ID ABIL_RAT STANDARD; PRT; 1321 AA.
 AC 070127;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bile salt export pump (ABP-binding cassette, sub-family B, member 11)
 DE (sister of P-glycoprotein).
 GN ABCB11 OR BSEP OR SPGP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=98212048; PubMed=9545351;
 RA Gerloff T., Stieger B., Hagenbuch B., Madon J., Landmann L., Roth J.,
 RA Hofmann A.F., Meier P.J.;
 RT "The sister of P-glycoprotein represents the canalicular bile salt
 RT export pump of mammalian liver".
 RL J. Biol. Chem. 273:10046-10050(1998).
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
 CC INTO THE CANALICULUS OF HEPATOCYTES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
 CC SITU.
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
 CC ATP BINDING CASSETTE (ABC) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U69487; AAC40084.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABC_tranpflrm.
 DR Pfam; PF00005; ABC_tran; 2.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Prodom; PD000006; ABC_transport; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transmembrane; Transport.
 KW DOMAIN 1 62
 FT DOMAIN 1 62
 FT TRANSMEM 63 83
 FT DOMAIN 84 147
 FT TRANSMEM 148 168
 FT DOMAIN 169 215
 FT TRANSMEM 216 236
 FT DOMAIN 237 240
 FT TRANSMEM 241 261
 FT DOMAIN 262 319
 FT TRANSMEM 320 340
 FT DOMAIN 341 353
 FT TRANSMEM 354 374
 FT DOMAIN 375 755
 FT TRANSMEM 756 776
 FT POTENTIAL.
 FT EXTRACELLULAR (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT POTENTIAL.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT POTENTIAL.

```

FT DOMAIN 777 794 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 795 815 POTENTIAL.
FT DOMAIN 816 869 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 870 890 POTENTIAL.
FT TRANSSEM 891 911 POTENTIAL.
FT DOMAIN 912 979 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 980 1000 POTENTIAL.
FT DOMAIN 1001 1011 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 1012 1032 POTENTIAL.
FT DOMAIN 1033 1321 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 455 462 APP (POTENTIAL).
FT NP_BIND 1113 1120 APP (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1321 AA; 146257 MW; 5443E4EF7B9FB1F6 CMC64;

Query Match 33.2%; Score 2230.5; DB 1; Length 1321;
Best Local Similarity 37.1%; Pred. No. 8.9e-124;
Matches 496; Conservative 248; Mismatches 490; Indels 103; Gaps 15;

QY 55 DKHRRKSSSSNNNAVSEVDALIAHLPEDEOVLTQLEIKVNTSFEGLMRYATKMDI 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 EENHAEESDSNH-----NDKSRLODKMEGDIRVGFELFESSKDI 59

QY 115 LIAVISTICAIAAASSTFORIML-----YQISYDE-----F 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 WLTLMGVCALTLHGMAOPGILIFIGINTDIFIKYDIERGLEIPKACVNTIWINSSF 119

QY 145 YDLT-----KNVYF--YVIGGEPTV--YVSTVGFIYGEHATOKIREYTL 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 HONMTGVTGCLVDISEMKFSGIYAGVMTVLIGYFOIRLMVITGARQIRMRKIYF 179

QY 190 ESILRONIGYFDKIGAEVTTTRITADPNLIDGISEKVGTLTALAFVTAFTIAYKYV 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 RRIIMREIGFDTGVGELNSRFADIEKINDIAIQAFLQRMSTAMCGLLIGFRGW 239

QY 250 KIALICSTVALVLTWGGSSQF---ITKYSKSLDSYGAGTVAEVISSIRNATAFG 305
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KLTLY-----ILAVSPLIGAAVIGLSIAKFTLELKAAYKAGSIADDEVLSIRTVAFG 295

QY 306 TQDLAKOYEVHIDEAKNCTKNOIYMGFMIGMFCMYSNYGLGFMGSRFLVD- GAVD 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 GENKEVERTEKNLVFAQRGIMKGMWGFPTGYMCLIFCYVALAFWYGSTVLDEEYF 355

QY 365 VGDILTVLAILISFSLGNVSPNAQFTNAVAAAKIFGTIDROSPLDPSYNEGKTLID 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 PGLTVGIFLCVILAAANIGHASSCLEIFSTGCAATINIFOTIDROPYIDCMGSDGYKLD 415

QY 425 FEGHIELRVNKHLYPSRPREVTVMEDVSLSNPAGKTALVPGSGKSTVVGIVEREYMPV 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 416 IKGEIFEHVNTFHYPSRPPVKIIDNLSMWIKPETTALVSSGAGKSTALQILQRFYDPC 475

QY 485 RGVYLLDGDHDKDLNRLRQQLSYVSQPVLEGTIIKRIHGLIGTKENSEDKVRE 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 EGVAVTLGDHDKDLNRLRQQLSYVSQPVLEGTIIKRIHGLIGTKENSEDKVRE 526

QY 545 LIENAKMANADFTALDEGYETNGORGEFLISGOKORIALARAVSDPKILLDEAT 604
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 527 DIVQAQKADNANVFTALQOQFPTLVGEGGQSGQKORVALARALINPKILLDMAT 586

QY 605 SALDTRKSEGVQAALERRAEGRTTIVIAHRLSTIKTAHNIIVLVNGKIAEQTHDELVD 664
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 587 SALLDNSEARVOALNKIQHGHITISVAHRLSTVRAADVIGFEGHVAVERGTHEELER 646

QY 665 GGAIRKLY--EAORINEQEA-----DALEDADAE-----DLTNADIAKITASSA 708
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 647 KGVYFVLVLOSQGDNAHETSIMGDATEGGLTEFTFSRGSYRDSLIRASIRQSKSOLS 706

QY 709 SSDLDCKPTTIRGTGTHKSVSASAILSKRPETPKYSIMTLTKFVASPNRPEIPVMLIGL 768
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 707 LITHDPEPLAVADHKSSYKDSKNDVLYEEVEPAP-----VARILKYNIPENWYILVGS 759

```

```

QY 769 VFSVLAGGQPTQAVILYAKAISTLSLPESOYSKRLHADFWSLMFVYGIQFTOSTNG 828
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 760 LSAINAAVPIYSLFESQLGTFSLDKF--QORSIHSMCLFVYLGCVISFTQPLQG 817

QY 829 AAFVCSERLIRARSTAFETILRQDIAFDKEENSTGALTSTETKHLGVSQVTLG 888
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 818 YTFKSGELLTKRLKRFKGFAMLGQDGMFDDLENNNGVLTTRLATDASQVATGSQV 877

QY 889 TILMTSTLGAIIIALAIAIMKIALVCSIVVPVLACGFRTFVYLAOFQSRKATYGS 948
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 878 MAAVSTNITLALIAAFSSWKSLLITTFPFLASGAQTKMLGFAQDQALERAG 937

QY 949 NFACATSSIRTYASLTREDDWEIYHAQDAGRTSLISVLSLLYASSQALVFECVA 1008
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 938 QITSEALSNIRTYAGIGVEGRFKAPEVELOTSKTAVRKANITYGLOFASQGIATFANS 997

QY 1009 LGFVYGTLLGHHEYDIFRFVCSSEILFQAQSNAGTVFSAPDMGAKNAAEFRLLDR 1068
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 998 AAVRYGYLLIAYEGELFSGSHVFRVVSVALSATAVGRTFSYTPSAKAKISAAREFOLLDR 1057

QY 1069 KPOIDMSEGEKLEIYEGELEFPANVHFRYPTREOPVLTGLDLYKPGQVALYSGC 1128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1058 KPLINYSAGEKMDNFQKIDFLDCKFTYPSRDQVNLGLSVNPGGTLAFVSGSGC 1117

QY 1129 GKSTTALLREFYDALAGSILVDGKDISKLINSYRSLSLVSGEPTLYGTTIKENTILG 1188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1118 GKSTQILDERFDPDQGTVMIDGHSKKVNIQFLRNSIGIVGEPLVFPDGSIMDNIKYG 1177

QY 1189 IYEDDVEEFLIACKDANIYDFIMSLPEGFNTVYSGKGMISGGQKORVALARALIND 1248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1178 DMTKEISVERAIAAKQAOLHDFVMSLPEKYEFTNVGQISQLSRGEKORAIARALVDRP 1237

QY 1249 KIILLDEATSLDSESEKVVQAALDAARGRTIIVAHRLSTIOKADVIVFPQOKIVES 1308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1238 KIILLDEATSLDSESEKTYQIALDKARBSGRTCIIVAHRLSTIONSIDIIVASQGVIEK 1297

QY 1309 GTHSELVQKRGRIYELV 1325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1298 GTHBKIMAGKAYYKIV 1314

```

Search completed: April 1, 2003, 16:01:16
 Job time : 38 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:57:13 ; Search time 57 Seconds
(without alignments)
4822.225 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLSTNPLSPETAMKEPAE.....YOKGRYELVNLQSLGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6704	99.7	1348	3	09Y8G1 emericella
2	5407.5	80.5	1349	3	043121 aspergillus
3	4575	68.1	1331	3	09HGT5 trichophyto
4	4167.5	62.0	1353	3	08X1F2 venturia in
5	3992	59.4	1307	3	043138 aspergillus
6	2682.5	39.9	1408	3	043140 cryptococcus
7	2519	37.5	1288	13	093437 gallus galli
8	2501.5	37.2	1287	13	091586 xenopus lae
9	2496.5	37.1	1272	11	091K64 ratius norv
10	2486.5	37.0	1272	11	0924L6 canis fami
11	2482	36.9	1280	6	046605 canis fami
12	2481.5	36.9	1285	6	002793 cris arles
13	2435.5	36.2	1169	11	060502 cricetus
14	2417.5	36.0	1275	11	08R427 ratius norv
15	2273	33.8	1348	13	090235 raja erinac
16	2272.5	33.8	1163	6	09T502 felis silve

17	2272.5	33.8	1407	5	08T9W5 dictyostell
18	2243.5	33.4	1293	3	09Y748 emericella
19	2232.5	33.2	1321	11	088331 ratius norv
20	2226	33.1	1284	3	09Y8G2 emericella
21	2224.5	33.1	1285	10	09LGX1 oryza sativ
22	2222.5	33.1	1432	5	08T9W4 dictyostell
23	2210.5	32.9	1289	10	0941H6 coplis japo
24	2205.5	32.8	1248	10	09FHE1 arabidopsi
25	2203.5	32.8	1278	10	09FWX7 arabidopsi
26	2200.5	32.7	1292	3	09C163 rhizomucor
27	2181.5	32.5	1286	10	080725 arabidopsi
28	2173.5	32.3	1275	5	061301 haemochus
29	2172	32.3	1262	10	08RVT7 triticum ae
30	2166.5	32.2	1292	10	09M1O9 arabidopsi
31	2154.5	32.1	1313	5	09V6Z6 drosophila
32	2153	32.0	1265	5	001495 caenorhabd
33	2148.5	32.0	1287	10	08S035 oryza sativ
34	2146	31.9	1230	10	09STI3 arabidopsi
35	2141	31.9	1313	3	08X1F3 venturia in
36	2135	31.8	1294	5	045721 caenorhabd
37	2132	31.7	1229	10	09FWX8 arabidopsi
38	2128	31.7	1229	10	049749 arabidopsi
39	2128	31.7	1229	10	09STI2 arabidopsi
40	2118.5	31.5	1240	10	09LHD1 arabidopsi
41	2106.5	31.3	1323	10	09MOM2 arabidopsi
42	2105	31.3	1302	5	024851 entamoeba h
43	2098.5	31.2	1310	5	024852 entamoeba h
44	2091.5	31.1	1283	5	024393 drosophila
45	2087	31.1	1252	10	09LXO arabidopsi

ALIGNMENTS

RESULT 1
ID 09Y8G1 PREDIMINARY; PRT; 1348 AA.
AC 09Y8G1;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Multidrug resistance protein MDR (ABC-transporter).
GN MTRD OR ABCD.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxId=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M6-096;
RA Andrade A.;
RT "Two novel ABC transporters from Emericella nidulans.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Nascimento A.M., Terenzi M.F., Goldman M.H.S., Goldman G.H.;
RT "Molecular characterization of ABC-transporter encoding genes in Aspergillus nidulans.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF071411; AAC43626.1;
DR EMBL: AF173826; AAF29805.1;
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABCtransporter.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR Prodom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
SQ SEQUENCE 1348 AA; 147467 MW; 780506AB631D0218 CRC64;

Query Match 99.7%; Score 6704; DB 3; Length 1348;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1334; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

QY 1 MSPLNPLSPETAMEPEPSTTEEOASTPHADEKILLSDSAPSSSTTATPADKEHRP 60
 Db 1 MSPLNPLSPETAMEPEPSTTEEOASTPHADEKILLSDSAPSSSTTATPADKEHRP 60
 QY 61 KSSSSNNAVSNEVDALIAHLPEDEROYLKTOLKEIKVNIISFGMLRYATKMDILIMVIS 120
 Db 61 KSSSSNNAVSNEVDALIAHLPEDEROYLKTOLKEIKVNIISFGMLRYATKMDILIMVIS 120
 QY 121 TICAIAA-----ASTFORIMLYOISYDEFDLTKNLYFYVYLGIGEPVY 166
 Db 121 TICAIAAGALPLFTLLFGLSLASTFORIMLYOISYDEFDLTKNLYFYVYLGIGEPVY 180
 QY 167 YSTVGEITGHAHOKIREYVLESTLRONIGYEDKLAGEVTRTADPNLQODISEK 226
 Db 181 YSTVGEITGHAHOKIREYVLESTLRONIGYEDKLAGEVTRTADPNLQODISEK 240
 QY 227 VGLTTLALTFVAFIAYVKKYKLLALICSSITVALVLPWGSGOFLIKYKSKLDSYGA 286
 Db 241 VGLTTLALTFVAFIAYVKKYKLLALICSSITVALVLPWGSGOFLIKYKSKLDSYGA 300
 QY 287 GGTVAEEVTSIRNATAFGTODKLAQOYEHVHDEAEKWTGKNQYNGFMIGAMGLMYSN 346
 Db 301 GGTVAEEVTSIRNATAFGTODKLAQOYEHVHDEAEKWTGKNQYNGFMIGAMGLMYSN 360
 QY 347 YGIGFMGSRFLVDGAVDGDILTVLMAITLIGSESLGANSPPNQAOTNVAANAALFEGT 406
 Db 361 YGIGFMGSRFLVDGAVDGDILTVLMAITLIGSESLGANSPPNQAOTNVAANAALFEGT 420
 QY 407 DRQSPDPYSNBEKTLDEHGEIENKRNKHIYSPREVTVMEDVSLMPAGKTALVGS 466
 Db 421 DRQSPDPYSNBEKTLDEHGEIENKRNKHIYSPREVTVMEDVSLMPAGKTALVGS 480
 QY 467 GSGKSTVGLVERFYMPVPGTYLLDGHDKDLMLRMLRQOISLVSOEPLYFGTTIKNIR 526
 Db 481 GSGKSTVGLVERFYMPVPGTYLLDGHDKDLMLRMLRQOISLVSOEPLYFGTTIKNIR 540
 QY 527 HGLIGKYENESBDKRELIENAKNANAHDTALPEGETVNGRGFLISGQORLA 586
 Db 541 HGLIGKYENESBDKRELIENAKNANAHDTALPEGETVNGRGFLISGQORLA 600
 QY 587 IAAVVSDEKILLDEATSAIDPTKSEGVYQAALERAEGRTTIVIAHRSTITANIVY 646
 Db 601 IAAVVSDEKILLDEATSAIDPTKSEGVYQAALERAEGRTTIVIAHRSTITANIVY 660
 QY 647 LVNGKTAEGTHDELVDREGAVRKLEVAORINEQKADALADADLTINADIAKIKTAS 706
 Db 661 LVNGKTAEGTHDELVDREGAVRKLEVAORINEQKADALADADLTINADIAKIKTAS 720
 QY 707 SASDDGKPTTIDRTGTHKSVSAILSKRPPTTKYSIMTLKRVASNREIYMYL 766
 Db 721 SASDDGKPTTIDRTGTHKSVSAILSKRPPTTKYSIMTLKRVASNREIYMYL 780
 QY 767 GLVFSVLAGGQPTQAVLYAKAISTLSPESQYSKLRHADWLSLAFVVGIIQFTQST 826
 Db 781 GLVFSVLAGGQPTQAVLYAKAISTLSPESQYSKLRHADWLSLAFVVGIIQFTQST 840
 QY 827 NGAAPVCSERLIRARSTAFRTILKODIAFDKKEENSGALTSFLSTETKHLISGVGYT 886
 Db 841 NGAAPVCSERLIRARSTAFRTILKODIAFDKKEENSGALTSFLSTETKHLISGVGYT 900
 QY 887 LGTIIMTSTTLGAIIIAIAIGKALVCIISVYPVLLAGGFYFYLAQFQSKLAYS 946
 Db 901 LGTIIMTSTTLGAIIIAIAIGKALVCIISVYPVLLAGGFYFYLAQFQSKLAYS 960
 QY 947 SANFACEASTSIRTVASLTRENDVWEIYHAODAGRTSLISVRSLLIYASQALVFC 1006
 Db 961 SANFACEASTSIRTVASLTRENDVWEIYHAODAGRTSLISVRSLLIYASQALVFC 1020

QY 1007 VALGFYGGTLLGHEDYFRFVFCSEILFGAQSAGTVFSFAPDMGKAKNAAEFRRLF 1066
 Db 1021 VALGFYGGTLLGHEDYFRFVFCSEILFGAQSAGTVFSFAPDMGKAKNAAEFRRLF 1080
 QY 1067 DRKPOIDNNSEEGEKLETVGELEFRNVHRYPTRPOPLRGDLTVPRGOVALVGS 1126
 Db 1081 DRKPOIDNNSEEGEKLETVGELEFRNVHRYPTRPOPLRGDLTVPRGOVALVGS 1140
 QY 1127 GCGSTTIALLEFPYALIGSILVGDKDISKLNINSRSLISVSEPTLYOCTIRENLI 1186
 Db 1141 GCGSTTIALLEFPYALIGSILVGDKDISKLNINSRSLISVSEPTLYOCTIRENLI 1200
 QY 1187 LGTIVEDVPEEPLIKACKKANITDFIMSLDEGNVYVSGKGLSGGQORVAIARALLR 1246
 Db 1201 LGTIVEDVPEEPLIKACKKANITDFIMSLDEGNVYVSGKGLSGGQORVAIARALLR 1260
 QY 1247 DPKILLDEATSAIDSESEKVVQALDAARGTTTAAVHRSTIOKAVIYVEDGKITV 1306
 Db 1261 DPKILLDEATSAIDSESEKVVQALDAARGTTTAAVHRSTIOKAVIYVEDGKITV 1320
 QY 1307 ESCTHSELVQKGRYVELVNLQSLGKH 1334
 Db 1321 ESCTHSELVQKGRYVELVNLQSLGKH 1348

RESULT 2
 043121 PRELIMINARY; PRT; 1349 AA.
 AC 043121;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Multidrug resistance protein 1.
 GN MDR1.
 OS Aspergillus fumigatus (Sartorya fumigata).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
 OC Eucotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5085;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=10AF/86/10;
 RX MEDLINE=98038972; PubMed=9373135;
 RA Tobin M.B., Peery R.B., Skatrud P.L.;
 RT "Genes encoding multiple drug resistance-like proteins in Aspergillus
 fumigatus and Aspergillus flavus.";
 RL Gene 200:11-23(1997).
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: U62934; AAB8658.1; -;
 DR EMBL: U62933; AAB8657.1; -;
 DR HSSP: P13569; INBD.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABCtransport.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00664; ABC_membrane_2.
 DR Pfam: PF00605; ABC_tran.
 DR ProDom: PD000006; ABC_transportr; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW ATP-binding; Transport.
 SQ SEQUENCE 1349 AA; 147723 MW; B90A86FD62C4165 CRC64;

Query Match 80.5%; Score 5407.5; DB 3; Length 1349;
 Best Local Similarity 79.1%; Pred. No. 2; 1e-310;
 Matches 1069; Conservative 114; Mismatches 141; Indels 27; Gaps 6;

QY 5 EYNPLSPETAMEPEPSTTEEOASTPHADEKILLSDSAPSSSTTATPADKEHRKSS 64
 Db 5 EYNPLSPETAMEPEPSTTEEOASTPHADEKILLSDSAPSSSTTATPADKEHRKSS 64
 QY 5 EYGAASSREKSL-EDLVATLEKGRSTSSGADNEKPHDHSISDITMAPPD---GKKRD 59
 Db 5 EYGAASSREKSL-EDLVATLEKGRSTSSGADNEKPHDHSISDITMAPPD---GKKRD 59
 QY 65 SNNAVSVNEVDALIAHLPEDEROYLKTOLKEIKVNIISFGMLRYATKMDILIMVIS 124
 Db 65 SNNAVSVNEVDALIAHLPEDEROYLKTOLKEIKVNIISFGMLRYATKMDILIMVIS 124
 QY 60 HGRAVDLND-DSLFAHLQHEKEVLRKQDAPSVKVSFTLYKYSARKDILLIIVSAICA 118
 Db 60 HGRAVDLND-DSLFAHLQHEKEVLRKQDAPSVKVSFTLYKYSARKDILLIIVSAICA 118

QY	125	IAA-----	-----	ASIFQOIMLYQISYDEDFYELIKRNVLYEVLIGEFVYVYST	170
Db	119	IAAGAAALPFTLIEFGSLASAPFOGSLCTMPYHEFYHFKLIRKNULYFYLGIAEFVYVYST			178
QY	171	VGFYITGEHAFOKIREYLYESILRONIGFYFDKIGAGEVTTRTITADTNLIQDGISERKGYLT			230
Db	179	VGFYITGEHLOKIRENYLTALIRROMNAPYDEKIGAGEVTTRTITADTNLIQDAISEKGYLT			238
QY	231	LTALATVTAFTIAYVYKMLALICSTYVALVLMNGGSSOFLIKSKSLDYSGAGTV			290
Db	239	LTAATVTAFTIAYVYKMLALICSTYVALVLMVNGGSRFLVYKSKSIESYSGAGTV			298
QY	291	AEEYISSIRNATAFGTODKAKOYEHLDAEAKWGNKNOIVNGFMGAMFGLMSNYGIG			350
Db	299	AEEYISSIRNATAFGTODKAKOYETHLDAEAKWGNKNOIVLGMGFMGFMFSNYGIG			358
QY	351	FWMGSRFLVDGADVGDILTVALMALLIGFSFLGNVSPNQAFTNAYVAAAKLFTGYIDROS			410
Db	359	FWMGSRFLVYGEVNGVGLVFLMSLIGFSFLGNVAPNQAFTNGVAAAKLYSTIDRS			418
QY	411	PLDYPISNEGKTLDHFEGHIELRNVKHIYSPRPYVYMEDVSI SMPAGKTALVPSGSGK			470
Db	419	PLDYPISOBGKVLDFEGENIEFRNVKHIYSPRPVYVYMEDVSI SMPAGKTALVPSGSGK			478
QY	471	STVYGLVERFMPARGVVLVDGDHIDIMLRMLROOISLYOSEPVLYFGTITKNTIRHGLI			530
Db	479	STVYGLVERFRLPYGGVGLVDGDHIDQTLRMLRROOISLYOSQPVLYFTTIRNIEHGLI			538
QY	531	GTKYENESDKVRLEIENAAKMANADFEITALPEGYETVNGVORGFLISGSGQKORALARA			590
Db	539	GTFKFEHESDKIRIELVENAARMANADFEITALPEGYDTNVGORGFLISGSGQKORALARA			598
QY	591	VYSPRKLILDDERTSALDPTSEGVQAALERRAEGRTTYIAHRSTIKTANHIVLVNG			650
Db	599	IVSOPKLILDEATISALDPTSEGVQAALDKAEGRTTYIAHRSTIKTANHIVAMVG			658
QY	651	KIADQGHDELDPBGAGYRKLVEAORINBEKDALE---DADAEDLTNAADIAKIKTASS			707
Db	659	KIADQGHDELVDPRKGYLYLVAQRINBEKEKALEADADADDDPGEQGYTRKTAIVS			718
QY	708	ASSDLDG---KPTTIDRTGTHKSVSSALISKRNPETTPKYSIMWTLTKVASFNPETIY			763
Db	719	SSNSLDVADEKARLEMKRKTGTQGSVSAYLSKKVPQEFQKYSIMWTLVYKIGAFNNPELGY			778
QY	764	MLIGLIVSVYLAGGQPOAVLYAKAISTYSLPESQYSKLIRHADDEWSLMFEVYGLIQPT			823
Db	779	MLIGLITFSPYLAGGQPOAVLYAKAISTYSLPESMFHKLRHANFWSLMFEVYGLIQFTS			838
QY	824	QSTNGAFAVCESELTBRASTAFRIILNODIAFPEKXENSGALTSFSTETKHLGYS			883
Db	839	LSTNGTAFALICESELTBRASQAFRSILNODISFPEKXENSGALTSFSTETKHLGYS			898
QY	884	GVTLGITLIMTSTYGAAILIILAIGKMLAVCISVVPVYLACGFYFYMIAQFQSRKLA			943
Db	899	GVTLGITLIMTSTYGAAMITIALIIGKMLAVCISVPIILACGLFRFYMLAQFQSRKSA			958
QY	944	YEGSASFACATSSIRTVASLTFRBDYWEIYHAQDDAGRTSLISYLRSSLLYASSQALY			1003
Db	959	YEGSASFACATSSIRTVASLTFRBDYWEIYHAQDDAGRTSLISYLRSSLLYASSQALY			1018
QY	1004	FEFCVALDFWYGGTILGHHEHIDIRFRVYVCSSEILFGAQSAGTVSPRPDMGKKNMAAEPK			1063
Db	1019	FEFCVALDFWYGGTILGHHEHSIRFRVYVCSSEILFGAQSAGTVSPRPDMGKKNMAAEPK			1078
QY	1064	RLFPFRKPOIDNMWSSEGEKLTVEGELEFRNVHRRYTPREOPVLRKGLDLYTVKPGYVALY			1122
Db	1079	KLFPFRKPTIOWSDEGSKLESMGELEFRNVHRRYTPREOPVLRKGLNLSYVPRGYITALY			1133
QY	1124	GPSGCGSTTIALIERFYDAIAGSLVADGKDISKLININSYRSLSLVSOEPTLYOGTIKE			1183
Db	1139	GPSGCGSTTIALIERFYDYDALAGVDFGDKITKLWNVNSYRSLSLVSOEPTLYOGTIKE			1198
QY	1184	NILGLYEDDVPREFELKACKDANIYDINSLPEGNNTYVVGSKGMLSGGOKORVALARA			1244

Dd	1199	NILGVDADVDSSEELLIKYCKANDATIDYDMSLPEGFDDIYVGKGMLSGGQXORVAIARA	1258
Qy	1244	LIRDRKLILLDEATSALDSESEKVVQAALDAARRTTIAVAHRLSTIQKADVIYFDDG	1303
Dd	1259	LLRDRKVLLLDDEATSAIDSESEKVVQAALDAARGRTTIAVAHRLSTIQNADIIVYFDG	1318
Qy	1304	KIVESGTSELVOKKGRIYYELYNLOSLGKH	1334
Dd	1319	KIVESGTHEHLIRNKGRYYELYNLOSLGKTH	1349
 RESULT 3 09HGTS PRELIMINARY; PR7; 1331 AA.			
AC	09HGTS		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Multidrug resistance protein MDR.		
GN	ATRD.		
OS	Trichophyton rubrum.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;		
CC	Oxygenales; Arthrodermataceae; mitosporic Arthrodermataceae;		
CC	Trichophyton		
OX	NCBI_TaxID=5551;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Fachin A.L., Ferreira-Nozawa M.S., Maccheroni W. Jr.,		
RA	Martinez-Rossi N.M.;		
RT	"Identification and Characterization of atrd gene, a new multidrug		
RT	ATP-Binding Cassette Transporter of Trichophyton rubrum."		
RU	Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.		
CC	-1 - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL; AF291822; AAC01549.3; -		
DR	InterPro: IPR003593; AAA_Arpase.		
DR	Interpro: IPR002106; AtRNA_LigaseIT.		
DR	InterPro: IPR001140; ABCtransprtTM.		
DR	InterPro: IPR003439; ABC_transportr.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	Prodrom: PD000006; ABC_transportr; 2.		
DR	SMART: SM00382; AAA_2.		
DR	PROSITE; PS00339; AA_TRNA_LIGASE_I1.2; UNKNOWN_1.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
KW	ATP-binding; Transport.		
SO	SEQUENCE 1331 AA; 145283 MW; E35B7AFCE3043B60 CRC64;		
 Query Match 68.1%; Score 4575; DB 3; Length 1331; Best Local Similarity 67.9%; Pred. No. 3e-261; Matches 901; Conservative 183; Mismatches 223; Indels 20; Gaps 4;			
Qy	26	EQASPPHADEKKLLISLSDAPSSTTATPADREHRPK--SSSSNNAVSVNEVALTAHP	83
Dd	3	EYSEKPNQDDBSVKOENRNRPASSSSSTSDEKVAKGSNDTKFSPTEDDAQLAHP	62
Qy	84	DERQYLKTLEIEIKNISFFGLMRATYKMDLIINVTSTCAIAA-----AS	129
Dd	63	HERELLKQOLFIPDKAKHYGTLEFRATRNDMFLIAIVSLASIAAGALPLFTVLGSLAG	122
Qy	130	TFOQTMLQISTDEFYDELTKNVLYFYVILGIFEVTVYVSTGVITYGSHATQIREYLL	189
Dd	123	JFRDALTRIRTYDFENSILTNRSLTFYVYLGIQFILLYSVTGFIYVGEHIOTIKRAYL	182
Qy	190	ESILKNQINGPYKLAGETVTRITADTNLIDGISSEKYGLTLTALTAFPTAFIIAYKKV	249
Dd	183	HAILKNQNGFEPKLAGETVTRITADTNLIDGISSEKYGLTLTALTSTFSAPFIIGVRYW	242
Qy	250	KIALICSTIYALVLTMGCGSQFIILIKYSKSLSDSYCAGGTVAEEYISSIRNTAFGTODK	309
Dd	243	KIALICSTIYAMILVMYGISRFVYKSGMTLVYSTGEGGTVAEEYISSIRNTAFGTODK	302
Qy	310	LAKOTEVHLDEAEKMGTKNQIYMFGMIGAMEGLMVSNTGLGFWMGSRFLVGDADVDTL	369

```
DB 303 LARQYEVHLKEARKGRRIOMLIGIMFGSMALMNTSYGLGFMGRSLFVLGGTLDLSAIV 362
QY 370 TVLMALILSFGISGVNVPNAQAFTNAVAAAKITGTIDRQSPDLPY SNECKTLDHEGHT 429
DB 363 NILLAIVISFSGISGVNAPNTQAFAAISAGAKITFTIDRVSAIDPQSDDEDTIENEGTI 422
QY 430 ELBNVNHITPSREYVYVMEVSLSMRAGKTTALVPGSGSKSVYGLVEFYMPVKTUL 409
DB 423 EFGIHNHITPSREYVYVMEVSLSMRAGKTTALVPGSGSKSVYGLVEFYMPVKTUL 402
QY 490 LQGHDIKIDNLRLRQOISLVSOEPLVFGTITIKNIRHGLIGIKYENESDEKRELIENA 549
DB 483 LQGRDITKIDNLRLRQOISLVSOEPLVFGTITIKNIRHGLIGIKYENESDEKRELIENA 542
QY 550 AKMANAHDTTALPEGYETNVGORGFLISGGOKORIAIARAVYSDPKITLLDEATSLDPT 609
DB 543 AKMANAHDTTALPEGYETNVGORGFLISGGOKORIAIARAVYSDPKITLLDEATSLDPT 602
QY 610 KSEGVVQAAALERAABERTTIVIAHRLSTIKTANIVYLVNGKTAEGTIDELVDRGAVR 669
DB 603 KSEGVVQAAALERAABERTTIVIAHRLSTIKTANIVYLVNGKTAEGTIDELVDRGAVR 662
QY 670 KLYEAGRIEOKKADALADADA--DLTNADIAKIKITASASDDLGRKPTTIDRTGTHK 726
DB 663 QLYEAGRIEOKKADALADADA--DLTNADIAKIKITASASDDLGRKPTTIDRTGTHK 722
QY 727 SVSSAALS-KRPETTPKYSMLTWLKFVASFNRPETPYMLIGLVSVLAGGOGPTQAVLY 785
DB 723 SLSSVILSQKRSQEKETEYSLSGLTIRFIAGFNPERLIMCGFEFVLSGAGQPVQSVFE 782
QY 786 AKAISLISLPEQSYKLRHDADPWSIMFVYGIQTITOSTNAAAVCSERILIRARST 845
DB 783 AGITTLISLPEQSYKLRHDADPWSIMFVYGIQTITOSTNAAAVCSERILIRARST 842
QY 846 AFTTILRODIAEFDEKENSSTGALTSFSTETKHSVSGVGTLLTMTSTTGAATIAL 905
DB 843 SFRAMLRQOIAEFDEKENSSTGALTSFSTETKHSVSGVGTLLTMTSTTGAATIAL 902
QY 906 AIGWKIALVCISVYVYVLLACGYRYTMAOFQSRKSLAYEGSANFACEATSTIRVVASLT 965
DB 903 AIGWKIALVCISVYVYVLLACGYRYTMAOFQSRKSLAYEGSANFACEATSTIRVVASLT 962
QY 966 RERDWEITHAOLADNOGRISLISVYVYVLLACGYRYTMAOFQSRKSLAYEGSANFACEATSTIRVVASLT 1025
DB 963 RERDWEITHAOLADNOGRISLISVYVYVLLACGYRYTMAOFQSRKSLAYEGSANFACEATSTIRVVASLT 1022
QY 1026 FFEVYVSEILFQAQAGTGVSEFAPDMGRKAKNAAEFRRLDFDRKPOIDMMSGEKLETV 1085
DB 1023 FFEVYVSEILFQAQAGTGVSEFAPDMGRKAKNAAEFRRLDFDRKPOIDMMSGEKLETV 1082
QY 1086 EGEIERFNVHFRYPYRPEQVYVRLGDLVYVPGQYVALVPGSGCGKSTTIALVERFYDA 1145
DB 1083 EGEIERFNVHFRYPYRPEQVYVRLGDLVYVPGQYVALVPGSGCGKSTTIALVERFYDA 1142
QY 1146 GSLVNGKDISKINISYVSEFSLVSOEPLVFGTITIKNIRHGLIGIKYENESDEKRELIENA 1205
DB 1143 GSLVNGKDISKINISYVSEFSLVSOEPLVFGTITIKNIRHGLIGIKYENESDEKRELIENA 1202
QY 1206 ANIYDFIMSLPEGFNTVYVSGKGLSGGOKORIAIARALIDPKILLDEATSLDPT 1265
DB 1203 ANIYDFIMSLPEGFNTVYVSGKGLSGGOKORIAIARALIDPKILLDEATSLDPT 1262
QY 1266 KVVQAAALDAAGRTTIAVAHRLSTIQKADVIYVFDQKIVESGTHSELVQKKRYEYLV 1325
DB 1263 KVVQAAALDAAGRTTIAVAHRLSTIQKADVIYVFDQKIVESGTHSELVQKKRYEYLV 1322
QY 1326 NQOSLGR 1332
DB 1323 HMOSSLER 1329
RESULT 4
```

```
08XIF2
ID 08XIF2 PRELIMINARY; PRT: 1353 AA.
AC 08XIF2;
DT 01-MAR-2002 (Tremblrel, 20, Created)
DT 01-MAR-2002 (Tremblrel, 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE ATP-binding cassette transporter ABC4.
OS Venturia inaequalis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes; at Chaetothromycetes incertae sedis; Venturiaceae;
OC Venturia.
OX NCBI_TaxID=5025;
RN [1]
RP SEQUENCE FROM N.A.
RA Schnabel G., Schnabel E.L., Jones A.L.;
RT "Complete sequences of four ABC transporters from Venturia
inaequalis."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF375879; AAL57243.1; -.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR001440; ABCTransporterTM.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran.2.
DR Prodom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_2.
DR ATP-binding.
SQ
SEQUENCE 1353 AA; 149136 MW; 7C8E9093D723D2C CRC64;

Query Match 62.0%; Score 4167.5; DB 3; Length 1353;
Best Local Similarity 61.0%; Pred. No.3.5e-237;
Matches 834; Conservative 184; Mismatches 290; Indels 45; Gaps 6;

QY 5 ETRNPLSPETAMRPAETSTFEQASTPHADEKILSDLSAPSTYATPDKERKSSS 64
DB 19 KTRGSGSEFEIKDKNARQMGAPY-----GNDTPSGGKELKMKRL 64
QY 65 SNNASV-NEVDALIAHLPDERQVLTQLEIKVNIISFEGIMRYATKMDILTMVSTIC 123
DB 65 DSNVIEVKQNDNDPFRHLPHEQEILIRQTFIPDVAKGVFTLYRASRMDMAWMVSVIC 124
QY 124 AIAA-----ASTFORIMLYISIDEFDELTKVLYVYIGIEFPT 165
DB 125 SIVSGAMPIMTVVFGGLGFLFADYPRKN---ITYKQFSELSHPVLYFLYLAIGTFY 180
QY 166 VYVSTVGFITYGHNATQIKREYLYESILRONIGYFPIKLAGETTYRTADTNLIQDIS 225
DB 181 TYIMTVGFITYGRCGKIRERILKAMLRQNIAPFKLAGETTYRTADTNLIQDIS 240
QY 226 KVGILLTALATFYATIAVVKYWKLLALICSTIVALVLTMGGSQFTIKYKSLDSYG 285
DB 241 KFGILTALATFASAVIAFIKYKWLTLITLTSVPFATLVMGVSSFFVVRWTVRSQTEYA 300
QY 286 AGGTVAEEVYISIRNATATGTOPLAKQYEVHDEAEKNGTKQIYMGFMIGMFLMYS 345
DB 301 KGGTIAEEVYISIRNATATGTOPLAKQYEVHDEAEKNGTKQIYMGFMIGMFLMYS 360
QY 346 NYGLGFMWMSRFLVDADVVDGILITVMAILLGSFSLGVNVPNAQAFTNAVAAAIFET 405
DB 361 NYGLSFMWMSRFLVDADVVDGILITVMAILLGSFSLGVNVPNAQAFTNAVAAAIFET 420
QY 406 IDROSPLDPSYNEGTLDHEGHTIELRNKHIYPSREYVYVMEVSLSMRAGKTTALVGP 465
DB 421 IDRTSPMDPSYNEGTLDHEGHTIELRNKHIYPSREYVYVMEVSLSMRAGKTTALVGP 480
QY 466 SSGKSTVYVGLVERFYMPYRAGTYVLLDGHDKIDNLRLRQOISLVSOEPLVFGTITIKN 525
DB 481 SSGKSTVYVGLVERFYMPYRAGTYVLLDGHDKIDNLRLRQOISLVSOEPLVFGTITIKN 540
QY 526 RHGLIGTKYENESDEKRELIENAAKMANAHDTTALPEGYETNVGORGFLISGGOKOR 585
DB 541 RHGLIGTKYENESDEKRELIENAAKMANAHDTTALPEGYETNVGORGFLISGGOKOR 600
```


QY 586 ATRAVVSDPKILLDEATSDLTSEGVQALERAEGRTTIVAHRLSTIKTAHIV 645
 DB 601 ATRAVVSDPKILLDEATSDLTSEGVQALERAEGRTTIVAHRLSTIKTAHIV 660
 QY 646 VLVNGIAEGTHDELVDRCAGARKLYEAKRI-----NEOKEDALDEDAEDLTNADI 699
 DB 661 VMSRGAIVEGGRHSELERSAFENLVFAORIAEIKNDPEEVEIIEQVEDGOLHRA-- 718
 QY 700 AKIKTASASDSDGKPTTIDRTGTHKSVASAIKSRPEPTTPKSLMTLTKFASPNRP 759
 DB 719 ATNEKEPIDPDEDEPVGRKARKQSGKSSISVELKRGKTQTPETSLIQLLGAVWSNKT 778
 QY 760 EIRPYMLIGLVSVALGGOPQOAVLYAKAISTSLPESQYSKLHDADFWSLMFVYGI 819
 DB 779 EMFIMLIGFICSVIAGGNGVQAIFFAKAVSALSLPQAQAEALRSEINFWSLMYLMAGT 838
 QY 820 QFTTOSTNGAFAVCSERLIRARSTAFRTILKODIAFPFKEENSTALSTFSTETKHL 879
 DB 839 QILSNFGQAVAFGYSERLIRARDOFRHMLRODIEFFRENNAGSLSTFSTSTOL 898
 QY 880 SGVSGVTLGTLMTSTLGAIIITIALIGKLLALVCISVVPVLLAGCFYRPVMAQOSR 939
 DB 899 SGSGSTLGIITLLOYSTTLVVAATISLAIGKLLALVCISAIPLVLAGCFYRPVMAQAR 958
 QY 940 SKLAEGSANFAECASISRTVASLTREPDWEIYHAQDLAOGRTSLISVRSLSLYASS 999
 DB 959 TKAAVYNSAGFACEATSAIRTVASLTREPDVEFEKTHAQLEQAOKASLSVLSKLSALVAA 1018
 QY 1000 QALVFCVVALGFTVGGTILGHEYDIFRFVYCSSEILFGQASAGTVSPADPKAKANAA 1059
 DB 1019 QSFIFCIIAGFMYGGTILAKREYSOFQFVCFMSIIFGQASAGTISFAPDMKAKANAA 1078
 QY 1060 AEFRRLEDRKPOQIDNMBSEGEKLETVGEIEFRNVHFRYPRPQVPLRGDLTVKRGQY 1119
 DB 1079 AELQITLDRPKIDCSEBDEDRILSYGAYEPRDYHFRYPRPQVPLRGDLTVKRGQY 1138
 QY 1120 VALVPSGCGKSTTIALLEFRYDALAGSLVDKDISKLINSYRSLSVSQDEPTLYOG 1179
 DB 1139 VALVAGSAGCGKSTTIALLEFRYDPLSGGVYVDGQEVSKLVANDRYSLVALVSQDEPTLYOG 1198
 QY 1180 TIKENILIGVEDDYPREFLIKACKONATIDFIMSLPEGRNTYVGSAGMISGQOKORVA 1239
 DB 1199 TIKENILIGVEDDYPREFLIKACKONATIDFIMSLPEGRNTYVGSAGMISGQOKORVA 1258
 QY 1240 IARALLDPKILLDEATSDLTSEGVQALERAEGRTTIVAHRLSTIKTAHIV 1299
 DB 1259 IARALLDPKILLDEATSDLTSEGVQALERAEGRTTIVAHRLSTIKTAHIV 1318
 QY 1300 FDOGKIVSGTHSELVOKKGRYYELVNLQSLGK 1332
 DB 1319 IDGGRVVEGSHNYLTSKNGRYAELVNLQSLGK 1351

RESULT 5
 043138 PRELIMINARY; PRT; 1307 AA.
 ID 043138
 AC 043138;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Multidrug resistance protein 1.
 GN MDR1.
 OS Aspergillus flavus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5059;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 24133;
 RX MEDLINE=98038972; PubMed=9373135;
 RA Tobin M.B., Peery R.B., Skatud P.L.;
 "Genes encoding multiple drug resistance-like proteins in Aspergillus

RT fumigatus and Aspergillus flavus.";
 RL Gene 200;11-23(1997).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: U62932; AAB88656.1; -;
 DR EMBL: U62931; AAB88655.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABCtransporter.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00664; ABC_membrane_2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Prodom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA_2; TRANSPORTER; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport.
 SQ SEQUENCE 1307 AA; 143874 MW; 2F65C12CF2F5F3 CRC64;

Query Match 59.4%; Score 3992; DB 3; Length 1307;
 Best Local Similarity 60.2%; Pred. No. 7.7e-227;
 Matches 790; Conservative 211; Mismatches 244; Indels 68; Gaps 11;

QY 60 PKSSNNAAVSVNEVDALIAHLPEDEROVLTQLEIEIKVINSFGLMRYATKMDILNVI 119
 DB 13 PKSGTGTSTT-----GHSVSHAEEVDRLQHTPVSQIGFGITRYATRMVDAILFG 64
 QY 120 STICAIA-----ASTFORIMLYQISYDEYDELTKNLYFYLLIGERVT 165
 DB 65 SALAAIAGGALPLFTVLFGRITSTFQDIATHRTYDHFHHLKNVYFYLLGAEEVA 124
 QY 166 VYVSTVGFITGEHATQKIREYILESLIRONIGFPAKAGAEVTRITADNLIDGISE 225
 DB 125 IYLATVGIYGDHVVOQIIREYFOALIRONIAEFDLIGAEIYTRITADNLIDGISE 184
 QY 226 KVGTLTALATFTVFAFIYAAYKYMKALICSTYVALVTMGGSOPFIKYSKSLDSYG 285
 DB 185 KVGTLALGSLFTVFAFIYAAYKYMKALICASLALITMGGSSTLMIFSKALIEVG 244
 QY 286 AGTVAEEVSSIRNATAFGTDQKIAQYEVHLDAAEKMGTKNOIVMGFMIGAMEGLMY 345
 DB 245 RGASMAEDILDSIRVAFAFNAQETLARKYFESHLDABGPKMSKVFYAIMVAGLILCY 304
 QY 346 NYGLGFMWGRSFLYDGA--VVGDIIVLYMAIILGSPSLGVNSPRAQAFYNAVAAAKIF 403
 DB 305 NYGLGFMWGRSFLYDGA--VVGDIIVLYMAIILGSPSLGVNSPRAQAFYNAVAAAKIF 364
 QY 404 GTIDRQSLDPLDYNEGKILDFEGHIELRNYKHIYSPPEYTVMEVDSLSPAGKTALY 463
 DB 365 GTIDRQSLDPLDYNEGKILDFEGHIELRNYKHIYSPPEYTVMEVDSLSPAGKTALY 424
 QY 464 GPSGSGKSTVGLVERFYMPRGTVLLDGHDIKDLNLMRLROQISLVSQEPLFGTTIYK 523
 DB 425 GPSGSGKSTIISLERFYDPVAGTILMDGHDIQTLNLMRLROQISLVSQEPLFGTTIYK 484
 QY 524 NIRGLIGTKENESDEKVRLEIENAKMANAHDFITLAPGEGYTNVQGRFLISGCGQ 583
 DB 485 NIRGLIGTKENESDEKVRLEIENAKMANAHDFITLAPGEGYTNVQGRFLISGCGQ 542
 QY 584 RIATARAIVSDPKILLDEATSDLTSEGVQALERAEGRTTIVAHRLSTIKTAHIV 643
 DB 543 RIATARAIVSDPKILLDEATSDLTSEGVQALERAEGRTTIVAHRLSTIKTAHIV 602
 QY 644 IVLVNGIAEGTHDELVDRCAGARKLYEAKRIENQREADLEDAEDLTN----- 696
 DB 603 IIVLVNGIAEGTHDELVDRCAGARKLYEAKRIENQREADLEDAEDLTN----- 660
 QY 697 -----ADIAKITASASDLDGKPTTIDRTGTH-----KSYSAALSK-R 736
 DB 661 MTEFFDKDPDDESDIYS--LSDASDI-----GLHTEKORPVSRLSLTLM 708
 QY 737 PPEITPKYSLMTLTKFVASFNRPEIPLYLIGLVPSVALGGOPQOAVLYAKAISTSLP 796
 DB 709 QPVAEBAVSFTLTKFLASFRKPEPFLITGLCASITLGGIOPSAVYFAKAVSTLSLP 768
 QY 797 SQYSLRHADDFWSLMEFVVGIIQFTTOSTNGAFAVCSERLIRARSTAFRTILKODIA 856

Db	769	LEYPRLRDANFNCMLFMIGITVLSVYSGTFLFAVSSSEKVMYRARSQAFLVILQDIS	828
Qy	857	FEDEKEENSTGALSTSLSTETFHLSGVSVYLTGLTILMTSTIGAAIITAIAGKTLAVCI	916
Db	829	FEEDQENTTGLATLNLSTGLSTELTGLISGVTLTGLTIIVSNVNLVSLGVALYIGKTLAVCI	888
Qy	917	SVVPELTLACGCFREFYMLAQFOSRSKLAIEGSAFNACEATSSIRTVASITREDRVMEYLHA	976
Db	869	SAVPALLMCGVVRWVWMLERFQPRRAKKVAQESASSACEAASAIRVVELTMTETALQSTQA	948
Qy	977	QLDQAGTSLTSLVLRSSLTSLVSSQALVFVCALGFWTGTGLLGHHEYDIFREFVCFSEIL	103
Db	949	QLRRLKSDIDPIRYKSSLTSLVSSQALPEFCMALGFWYGGSLTGHGESELPQFYVCFSEVI	1000
Qy	1037	FGAAGAGVFSEFADPMGAKAAAEFRRLRPRKQDIDNMSEEGKLTVESELEFRVHF	109
Db	1009	FGAAGAGVFSEFADPMGAKAAAEFRKLTSSDTMHASRS-KGVPAVSMKGLVEFRVVF	106
Qy	1097	RYPRPEPQVLRGDLVLPKQOQYALVPSGCGSKSTTALLERFYDAIAGSLVDKDIS	115
Db	1068	RYPSRLQEPILRLNLTLTKPQFALVAGASSGKSTTALLERFYDPLKGVYVDGNII	112
Qy	1157	KLNTNYSFSLVSOEPTLYQGTFIKENILGLIVEDDVEEFILKACKDANIYDFINSLP	121
Db	1128	TLEMSYSKSHALLISOEPTLTQGTIRENIIILGSSNTPHYTDFTLAKACKDANIYDFILSLP	118
Qy	1217	EGFNTVYVSGKGMISGGCKOKVATARALTRPKITLLDDEASALDSSEKVVQAALDAAA	127
Db	1188	QGFNTVYVSGKGMISGGCKOKRIATARALTRPKITLLDDEASALDSSEKVVQAALDAAA	124
Qy	1277	RGRTTAAVNHLSLTOKRADVYIVYVDQKIVYESGTHSELVQKRGYELVNLDS	1329
Db	1248	RGRTTAAVNHLSLTQIRADLTLYVDQGVVBSGTHRELKRKGYYELVNLQN	130
RESULT 6			
ID	043140	PRELIMINARY;	PRT; 1408 AA.
AC	043140;		
DT	01-JUN-1998 (TREMblrel. 06, Created)		
DT	01-JUN-1998 (TREMblrel. 06, Last sequence update)		
DT	01-JUN-2002 (TREMblrel. 21, Last annotation update)		
DE	Multidrug resistance protein 1.		
GN	MDR1.		
OS	Cryptococcus neoformans (Filobasidiella neoformans).		
OC	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;		
OC	Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.		
OX	NCBI_TaxID=5207;		
RA	SEQUENCE FROM N.A.		
RC	STRAIN-M1-106;		
RX	MEDLINE=96072425; PubMed=9409767;		
DR	Thorneswell S.J., Peery R.B., Skatrud P.L.;		
RT	"Cloning and characterization of CEMDL: a Cryptococcus neoformans		
RT	gene encoding a protein related to multidrug resistance proteins.";		
RL	Gene 201:21-29(1997).		
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL; U62930; AAC48890.1; -;		
DR	EMBL; U62929; AAC4889.1; -;		
DR	InterPro: IPR003593; AAA_ATPase.		
DR	InterPro: IPR002106; AATRNA_ligaseII.		
DR	InterPro: IPR001140; ABCtransprtTM.		
DR	InterPro: IPR003439; ABC_transportr.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	ProDom; PD000006; ABC_transportr; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00339; AA_TRNA_LIGASE_II.2; UNKNOWN_1.		
DR	APR-headers; PS00211; ABC_TRANSPORTER; 2.		
DR	Anti-binding; Transport.		
QO	SEQUENCE 1408 AA; 152143 MW; CEBDA34A09A6F69F8 CRC64;		

Query Match	Similarity	39.9%	Score 2682.5	DB 3	Length 1408
Best Local	Similarity 43.9%		Pred. No. 1.9e+149		
Matches	597	Conservative 233	Mismatches 449	Indels 81	Gaps 16
QY	33	AADEKKILILSDAPSSTTA-TPADEEHRRKSS-SSSNNAVSNEYDALIAHLPEDEQVL	89		
Db	69	AAEGKISRSSIAASDILKNSPLEK---PLSNAFSKSHPKSKNDPLKSRKKKEEEK	125		
QY	90	KTQLE-----IKNIFPFGIMRYATKMDILINVISITCAIAAS-----T	130		
Db	126	NKEKERKASVLPPEVFFALFREAALEIIMVLGLVLAAMGSCPLMTLIFGRLLTSET	185		
QY	131	FORIMLYOISY-----DEFYDELFTNVLVYFVLGEGEPTVVSIVGFIY	175		
Db	186	NTAVIANIOSGGILPERSALQAAKDKDILKTOSGHNAITLMAIGMFLATLYMFIMV	245		
QY	176	TGEHATQKIREYLLSILRONIGYEDKLAGEHVTRITADNMLIDIGISEKGLVLTALA	235		
Db	246	TGELMSKRIREYRLAAVLRQETAYFPDDLQAGAVARIQDCLVQEGISEKALVFOYAG	305		
QY	236	TFVYAFITAYKMKALICSTYALVLMGGSGQFIITKSKSLDYSAGGVAAEVIY	295		
Db	306	TFVCFVYAFVAPSPRIAGALYSILEVIMLCGIMMTAAKYATAALDHIARAGSLAEVIY	365		
QY	296	SSIRNAFTPGTODKLAKOYEVLHDEAEKMGKTNOIYVGMIGAMGLMYSNGLEFPMKS	355		
Db	366	GISRTVOAFGKKILGDGFADHIEOSKIVYGRGSIFFEGGSLIMFFVYIAALAFYYGG	425		
QY	356	RFVYGAVDVGDILTVLMAILIGFSLGNVSPNAQFTNAVAAKIFCTIDROSPLDY	415		
Db	426	ILVSGQADSGIVINVFMSILIGSSMMMLAPDELAAYIKARAALKFATIDRYVAIDSA	485		
QY	416	SNEGKTLDFBQHIELRNVKHIIYPSRPEVYVMEVYSLSPAGKTALVGPSSGSKSYVG	475		
Db	486	SEEGKPRDGLRGEISFEFVWKFEYHPSPSIPILKGTTFTEAGKTPALVAGSGSGKSYVS	545		
QY	476	LYERYVMVRGVLLDGDIDKDLNRLRQOISLYSOEPLYFGTIIYKIRIGLLGTKE	535		
Db	546	LIERYDYPSGVVKIDGKDINSILNMLRQOIGLVSOEPLTFGTIVRGVGEHLIGSRE	605		
QY	536	NESEKVEKELLENAKMANAHDFITALPEGYETNGGKGLLSGGOKORIALARAVSPD	595		
Db	606	NASLEEKRELVKKACVDNANAHNFINKLQGYDTMNGEKMILLSGGOKORVALARAVSPD	665		
QY	556	KILLIDEATSLDPIKSEGVQOALERAEGRTTIVIAHRLSTIKYAHNITVLYNKAQD	655		
Db	666	RILLIDEATSLDPIKSEGVQOALERAEGRTTIVIAHRLSTIRADRIYVGGGEVLQD	725		
QY	656	GTHDELY-DRGAYKLYLEAQRINQKEDALE-DADAEGLDNLADIAIKTRASASSDLD	713		
Db	726	GSHNDILANENGPAAQVLVNNOKLAEAAEAALQVDDIEDPDA-----VFIGGSPMQ	779		
QY	714	GKPTTIDRTGTHKSVSA-----ILSKRPPEY-----PKYSIMTLL-----KFVASF	756		
Db	780	EKKOQLHNAVYGRSLASTIAMDDIOAKRAEYVAAGEKITSSRGYLRLLRMSADKFT---	836		
QY	757	NRPEIPLYMLIGLVSVIYAGGQPTQAVLYANAKISTLSPEOSYKILRHADAFWMSLMFFVY	816		
Db	837	-----YIAFIAICAGMYPPSLAILFGKALSOFELQD--ALRRALSSALMYIT	887		
QY	817	GIQITTIOSYNGAFAVCSERLIRARSTAFRTILKODIAFDKEENSTGALTSFLSTET	876		
Db	888	ALAAAFVIEFOSAGSRSGKWDINGLRRKLLFTATLRHDIEMVEDERNSTGAVTSMIDQP	947		
QY	877	KHLGVSQVTLTGILMTSTTGAAITILALIGWKALVACISVPPVLLACGFRTFMADQF	936		
Db	948	QKVQGLFEPITLGIYVQASCATLIGCCTIGLCYGPLALLIGLICIPILVSGGYIRLKVYAK	1007		
QY	937	OSRSKRLAEGSANFACETSSIRIYVATSLTREDYWEIYHAOLDAGRSLYSLSLSSLLY	996		
Db	1008	DQRMKMLHAASAHLSAAGAAYKIVYASLTREKDYARITSEALKAPMKLNFRSTKSQCLF	1067		
QY	997	ASSQALVEFYVALGFWYGGTLLGHHEYDIFRFVCFSEILFEGAOSAGTVSFADPMGRK	1056		

```

Db      1068 AASGGLFECILALVFYIGALMITIDAKYSTASFETVLNLSHFVASYIQAGNPFTFPDASKAN    112
Qy      1057 NAAAEFRRLFDKRPQIDNMWSESEKLE--TWGEIEIFRNVRHYEPTRRBPYLRIGLDTLV    111
Db      1128 SSAASIRSDINEPFIATIAESNEEKVDLHKHVVGHVRIEGVHEFYPPRPGCVRYLRNLITDY    118
Qy      1115 KPQGYVALVPSPCGSKSTTALLEREPDAIAGSLIVDGKDISKLNTNSRSFLSYLSQP     117
Db      1188 PAGTYIALVPSGGSGKTQTOMLERFYDLAAGVTLDGIDIKELINTASTYSQSISTSQEP     124
Qy      1175 TLVGCTKEKNLICIVE--DDVEPEEFLIKACRKANIYDFMTSLPEGFNTVSGSKGMLSG    123
Db      1248 TLYVGTRFNFLLCANKPIEEVYODEIDAACKRANIYDFVLSLPDEFDEYGVGKSQLSG    130
Qy      1233 GOKORVAIARALLDPKILLIDLATSALDSSEKKVQAALDAAGAARTTAIAHRLISTO    129
Db      1308 GOKORIAIARLIARNPKVPLLLIDEATSLDSOSEKVVQEALEDKAARKRTTIATHRLSIO    136
Qy      1293 KADVIYVEDOGKIYESTGESTHELVOKKGRYYELYVNOSLGK    1332
Db      1368 HSDRIYFSEGRAVHEGTQHOLLAKKGYYELVOMONLSR    1407

RESULT 7
Q93437 PRELIMINARY; PRT; 1288 AA.
DC      093437;
AC      093437;
DT      01-NOV-1998 (TREMBREL_08, Created)
DI      01-DEC-2001 (TREMBREL_19, Last sequence update)
DT      01-JUN-2002 (TREMBREL_21, Last annotation update)
DE      ABC transporter protein.
GN      CMDRL.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC      Gallus.
CX      NCBI_TaxID=9031.
[1]
RN      SEQUENCE FROM N.A.
RP      TISSUE=INTESTINE;
RX      Edelman H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C., Kane S.E., Kuchler K.; "Chdr1, a chicken p-glycoprotein, confers multidrug resistance and interacts with Estradiol."; Biol. Chem. 380:231-241(1999).
RC      -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR      EHBL_AJ009799; CA08835.1; -.
DR      HSHP_P13569; INBD.
DR      InterPro: IPR003593; AAA_ATPasE.
DR      InterPro: IPR001140; ABCTRanprtTM.
DR      InterPro: IPR003439; ABC_transportr.
DR      InterPro: IPR000504; RNA_rec_mot.
DR      Pfam: PF00664; ABC_membrane; 4.
DR      Pfam: PF00005; ABC_tran; 4.
DR      ProDom: PD000006; ABC_transporter; 2.
DR      SMART: SM00382; AAA; 4.
DR      PROSITE: PS00211; ABC_TRANSPORTER; 4.
DR      PROSITE: PS00030; RRM_RNP_1; UNKNOWN_2.
KW      ATP-binding; Transport.
FT      CHAIN         2       1288          ABC TRANSPORTER PROTEIN.
SC      SEQUENCE   1288 AA;  141917 MW;  CB258A5F2826DBC CRC64;

Query Match           37.5%; Score 2519; DB 13; Length 1288;
Best Local Similarity 41.2%; Pred. No. 7.2e+140;
Matches 554; Conservative 230; Mismatches 451; Indels 110; Gaps

Qy      32 HADEKITLSLASPSSTTTATPADKEHREPSSSSNNAVSVNEYDALIAHLPEDEROVLTKT    91
Db      2 HSEDK-----QRHTVGNGTYEIANMSODD-----PEDEKKGKKK    37
Qy      92 QLEBKNISSFGLMRRTAKMDILLNVISTICALAAASTFORIMLY-----        137
```

[illegible]

Db 1079 SSGCGSTVQVQLERFYPDIAGSEIVFPDDIDAKTINIQWLAISHIGIVSOEPILEPFTIAEN 1138
 QY 1185 ILIGIVEDYPEEFLLIACDANDIYFIMSLPEGEFNTVSGSGKMGSGGOKORVALARAL 1244
 Db 1139 IAGDGNREYVSHIELISNAASAISIFIDSLPEKYNTRVGDKGTLGGGOKRIAIARAL 1198
 QY 1245 LRDPKILLDEATSAIDSESEKVVQALDAARGRTIYVAHRISTIOKADVIYVPEQGR 1304
 Db 1199 IRKPQILLDEATSAIDSESEKVVQALDAARGRTIYVAHRISTIOKADVIYVPEQGR 1258
 QY 1305 IVESGTHSELVQKRGYELVNLQS 1329
 Db 1259 VIEGTHQQLLAKRGFYSLVNVQS 1283
 RESULT 8
 Q91586 PRELIMINARY; PRT: 1287 AA.
 AC Q91586:
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, last annotation update)
 DE Multidrug resistance protein.
 GN XEMDR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95322451; PubMed=7599185;
 RA Castillo G., Shen H.J., Horwitz S.B.;
 RT "A homologue of the mammalian multidrug resistance gene (mdr) is
 RT functionally expressed in the intestine of *Xenopus laevis*."
 RL Biochim. Biophys. Acta 1262:113-123(1995).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: U17608; AAA75000.1; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABC_transporter.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00664; ABC_membrane_2.
 DR Pfam: PF00005; ABC_tran_2.
 DR Prodom: PD000006; ABC_transporter_2.
 DR SMART: SM00382; AAA_2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport.
 KW SEQUENCE 1287 AA; 141505 MW; 06E95073C5771415 CRC64;
 SQ
 Query Match 37.2%; Score 2501.5; DB 13; Length 1287;
 Best Local Similarity 42.6%; Pred. No. 7.8e-139;
 Matches 545; Conservative 215; Mismatches 455; Indels 63; Gaps 12;

QY 366 GDILITVMAILIGSEFSICGNVSPNAQFTNAVAAAKIFGIIIDROSPLIDPYSGNEGTLDHF 425
 Db 339 GSYLVTFEFAVIGFAVAGQSPNIEAFANAGAAVYTFINIIDNPKIDSEKGLPKDKI 398
 QY 426 EGHLELNVAHYPSREVMEDVSLMPAGKTALVSGSGSKSVVGLVEFYFVR 485
 Db 399 KGDIENKNAVITYPSRKDIQVLKGLNLPISQKVALVSGSGSKSTVGLIQEFYDPE 458
 QY 486 GTVLLDHDIDKIDLRLMRLROOISLVSOEPLYFGTITKNIRHGLIGTKYENESDEKREL 545
 Db 459 GVTLDGQDRLSNIRLREIIGVSOEPILEPFTIADNIRY-----REDYVKEE 509
 QY 546 IENAKAANAHDFTALPEGETVNVGQRFLLSGGOKRIAIARAVYSDPKILLDEATS 605
 Db 510 IERATKEANAYDFIMKLPDLLETLYGERGTQLSGGOKRIAIARALVRNPKILLDEATS 569
 QY 606 ALDPTKSEGVQALERAERERTIYVAHRISTIKTANIVLVNKGIAEGTIDELVDRG 665
 Db 570 ALDPTSEAVVQSAIDKAREGRTIYVAHRISTIRNANAIGFQNGVTEGSHRELMERG 639
 QY 666 GAYRKIVLEAORINEQKEADALEDAE-----DLTNADIAKIKTASSASSDDGKRPETI 719
 Db 630 GYFENLVTLQFVETSKDTE--EDLETIYEKKIPYTHHSNLYRKSSRNITISKYPETE 687
 QY 720 DRTGTHKSVSSAILSKRPETTPKYSILMTLLKRVASEFNREIPYMLIGLVSVLAGGGP 779
 Db 688 D-----KEVDEE--EKKEEGPPVSPFVKM-----LNKREWFYFVGVICAMINGAT 726
 QY 780 TOAVLYAKAISTSLPESQSKLRHDAEFWSLFEVYGIQFTQTSNGAFAVCSRLI 839
 Db 737 AFALITFSRIIGVAPGPSQ--MSESSMYSLFLAIGVGFSTFPLQGTTFKAGEIIL 793
 QY 840 RRASTAFRTILRODIAFEPEKENSIGALTSFSTETKHLGSGVYIGTILMTSTTIGA 899
 Db 794 MRRLISFKMLQNGEIGFMDSKNSIGALTTRLATDSOVQAGTGRLLAALANVANT 853
 QY 900 AIIIALAIGKIALVCISSVPLVLAGCFYFYMALQFOSKILAYEGSANFACEATRSIR 959
 Db 854 AIIISFYGWQLTLILAIYPIVIAAGLVEMKMFAGHAKDKKELEKAGKISTDAVLNIR 913
 QY 960 TVASLFRERDWEIYHNOADQRTSLISVRSLSLYASSQALVFPYVALGFYVGGTL 1018
 Db 914 TVVSLFRERKFEAMETKSLGEPYRNISIKKALHGLTGLSOAHVVLCLCWFSVLCAVYL 973
 QY 1019 --GHHEYDIFRFVCESEILFGAQSAGTVSFAPDMGKANAAREFRLEFRKPDIDMS 1076
 Db 974 VEGILMKLD--EVFLVSAIYVIGAMALGQTSFAPDITKAMISAHAHIFSLIERYPQIDYS 1031
 QY 1077 EBEKLETVGELEFRNVHFRITRPEQVPLRGDLTVKFGQYVALVSGSGSKSTTIAL 1136
 Db 1032 DQGEKPKNCSGNVVFKGVNFNYPTRPDITVLOGLDISVKGETLALVSGSGSKSTTVSL 1091
 QY 1137 LEFFYDAIGSLVDGDKSLKINSYRSLISVSOEPILYOCTIENILIGIVEDYPE 1196
 Db 1092 LEFFYDPEGEVAVDGLSVANLNIQWVRADMGIVSOEPILEPFTIADNIRY-----REDYVKEE 1151
 QY 1197 EFLIKACKDANIYDITMSLPEGEFNTVSGSGKMGSGGOKORVALARALRDPKILLDEA 1256
 Db 1152 EEIETAKKANINHSFIESLTDKNTRYGDKGTOLSGGOKRIAIARALIRKPKILLDEA 1211
 QY 1257 TSAIDSESEKVVQALDAARGRTIYVAHRISTIOKADVIYVPEQGR 1316
 Db 1212 TSAIDSESEKVVQALDAARGRTIYVAHRISTIOKADVIYVPEQGR 1271
 QY 1317 KGRYVELVNLQSLGKH 1334
 Db 1272 LKGVFSLVTLQ--LGH 1286
 RESULT 9
 Q9JRK64 PRELIMINARY; PRT: 1272 AA.
 AC Q9JRK64:

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Multidrug resistance protein 1a.
 GN PGY1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSTAR; TISSUE=LIVER;
 RA Hootveld G.J.E.J., Wilms J.W.J., Hagenbuch B., Jansen P.L.M.,
 RA Meijer D.K.F., Muller M.;
 RT "Cloning and functional characterization of the rat multidrug
 RT resistance protein Mdr1a."
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: AF257746; AAF69007.1;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABCtransporter.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF000664; ABC_membrane_2.
 DR Pfam: PF000005; ABC_tran_2.
 DR ProDom: PD000006; ABC_transporter_2.
 DR SMART: SM00382; AAA_2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport.
 KW SEQUENCE 1272 AA; 140326 MW; B5C1DF89DFC986C5 CRC64;

Query Match 37.1%; Score 2496.5; DB 11; Length 1272;
 Best Local Similarity 42.2%; Pred. No. 1.5e-138;
 Matches 536; Conservative 225; Mismatches 460; Indels 49; Gaps 10;

QY 90 KTOLEIKVNISEFGRLKRVAKMDILIMVISTICAIASSTFGKIML----- 136
 DB 23 KKEKKKKPVAVLTMTMYGLMDREFYMLGTLAIHIGALPLMLVFGDMTDFANVG 82
 QY 137 -----YQIS--YDFEYDELTKNVLKLYVYLGIEFVTVVSTVGFIYTGKATQIKREY 187
 DB 83 NNSMSYNNMTDIYAKLEDEMTYAYYYTIGAGVLLVAVYQVSLMCLAGROQIHKIRK 142
 QY 188 YLESILKONIGYFDKLGAGEVYTRITADTNLIQDISKVGTLTALATVETAIYVK 247
 DB 143 FFRHINMOEIGMEFVHVGELNTRLTLDVSKINIGIDKIGMFQAMATFGFGLIGPTR 202
 QY 248 YKRLALICSSSTIYALVLTMGGSQFIITKYSKLSDSYGAGTVAEVISSIRNATAGTQ 307
 DB 203 GWKLTLLTILAIISPVLSAGIWAKLISFTDKLQAYAKAGAAVEYLAALRTVIAFGQ 262
 QY 308 DKLAQYEVHLDEAEKMGFTKNOIYMGFMIGAMFGIAMSNTGLGFMGSRFLVDCGAVDGD 367
 DB 263 KKELEKRNNNLEKAKRIGIKKAITANISMGAFLITIASYALATWYGSIVLSKEYITGQ 322
 QY 368 ILTVLMAILIGESFLGVNSPNAQATNAVAAAAKIFSTIDROSPLDPSYNEKGLDHEEG 427
 DB 323 VLVVFEVLIGAFSVGQASPIEAFAANRGAAYEVFSIIDNKPISDFSKSGHPRDNLG 382
 QY 428 HIELRNKHIYPSREPTVMEVDSLMPACKTALVGPSSGSKTVNGVLERFMPARGT 487
 DB 383 NLEKRNHFESYPSRKDVQILKGLKVKSGQVTLVNSGGKTRITVQLRLRLDEATIGE 442
 QY 488 VLLGDHDIKDLNRLMLRQOISIVSQEPVLFPTTYKNIHMLIGTKYENESDKVRELIE 547
 DB 443 VSIDGQDIRTINRYKYLEIIGVSVQEPVLFATTAENIRIG-----KENTYMD-----IE 493
 QY 548 NAAKMANAHDFITALPEGYETNGQRFLLSGGQKORITAIARAVYSDPKILLDEATATL 607
 DB 494 KAVYEAANAAYDFIMKLPHKFTLVGEBRGAQSLSGGQKORITAIARALVRPKILLDEATATL 553
 QY 608 DTSEGEVVOAALERAAGRTTIYAHRLSTIKTAHNTVVLVNGKIAAGTDELYDVGGA 667
 DB 554 DTSEEAVVQALDKARSGRTTIYAHRLSTVRNADVLAGFDGVIYVQGNHDELMREKGI 613

QY 668 YKRLVEAQRINEQEA--DALEDADAEIDLINADIKTATASSASDDLDGKPTTIDRTGTH 725
 DB 614 YKRLVMTQAGNELEIGNEACESKDGID--NVDS---SKNSGSLIPRSTRKSRINPH 668
 QY 726 KVSASAILSKRPETTPKYSILMTLKVFASPNRPEIPMLIGLVFSLVAGGQPTQAVLY 785
 DB 669 DQDELSTKREKLDADVPPASFWRLK---LNSTEMPVYVGVCAIINGLQAPASITF 724
 QY 786 AKAISTLS---LPESQSKLRHDADFMELMFVVGIIQFTQSTNGCAFAACSELRPRA 842
 DB 725 SKVGVFTKNDPTETQ-----RONSILSLFLILGILSFTTFIQTGFTGAGETILKRL 780
 QY 843 RSTAFRTILRODIAFFPKKEENSTGALTSFLSTERKHLGSGVTLTGITLMTSTLGAALI 902
 DB 781 RYVWFKSLRQDISMPDPPKNTGTLTRRLNDAQVAGAGSRILAVITQUNIANLGITII 840
 QY 903 IALIGKRLAIVCISVYVLLACGFYRPMLAQOSRKLAYEGSANPACATSSIRTV 962
 DB 841 ISLIYQWQLTLLALVPIIILAGVEMKMLSGQALKDKKELESGKLTATEAIEFRIVY 900
 QY 963 SLTERDQWEIYHQAQDAGRTSLISVRSLSLVAASQALVFCVAGLFWYGGTLGHNH 1022
 DB 901 SLTERQKFEIYTAQSLQIPYNNALKKAHVFGITTSFTQAMVFSYACFRGAYLVAREL 960
 QY 1023 YDIFREFESELIFGASAGTGVESFAPDMGKANAAEFRRLPDRKPOIDMSEEGEKL 1082
 DB 961 MFEFNVLLVFSAIYFGAANAQVGSFADYAKAKVASHIIRIETKTEIDSYSTEGIKP 1020
 QY 1083 ETVEGEIEFRNVHFRPTREPOVLRGLDLTVKQGYALVSPSGCKSTTALLERPYD 1142
 DB 1021 NMLEGNVAFKFNVMYFNPRPMIPVQLGSLTEVKGQTLALVSSGCKSTYVQOLLERYD 1080
 QY 1143 AIGSLIVDGKDISKLNINSTRSLISVSOEPTLYQGRIKNIILGLIVEDVPEERFLKA 1202
 DB 1081 PMAGTVFLDGKEIKQLANQMLRAHGLIVSQEPILFSDSIANLAVGDSRVVSHETVKA 1140
 QY 1203 CKDANIYDFIMSLEPEFNTVVGSKGMLSGQKORVATARALLDPKILLDEATATLSD 1262
 DB 1141 AKENIHIFQIDSLPEKVTNRGDKGTQSLSGQKORITAIARALVQPHILLDEATATLSD 1200
 QY 1263 ESEYVVOALDAAARGRTTIYAHRLSTIQKADVIYVFDGKIYESGTHSELVOKKGRY 1322
 DB 1201 ESEYVVOALDAAARGRTTIYAHRLSTIQKADVIYVQNGVYEHGHQOOLAKGIYF 1260
 QY 1323 ELVNLQSLGR 1332
 DB 1261 SMVSVOAGAK 1270

RESULT 10
 Q924L6 PRELIMINARY; PRT: 1272 AA.
 AC Q924L6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE P-glycoprotein.
 GN MDRLA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Yamazaki M., Ieake B.F., Kim R.B.;
 RT "Molecular Cloning of Rat Mdr1a cDNA."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF286167; AAK63023.1;
 DR InterPro: IPR001140; ABCtransporter.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF000664; ABC_membrane_2.


```

Db 203 TGFIVGFTPGWKLTLYLALISPVGLSAAIWAAILSSFTKELLAYAKAGAAVEVLAAI 262
OY 299 RNATJAGTODKLAKEYEHIDEAEKMGCTKNOIYVGMIGAMFGIMSNYNGIGFMWGSRLF 358
Db 263 RYVIAAGGOKKELEKRYKNKELEAKRIGIKAITANISIGAAFLITAYSTALAFYTGISLV 322
OY 359 VDGAVDVGDILTVLMAILIGSFISGVNPAQFTNAVAAAKIFGTIDRQSPLDPSNE 418
Db 323 LSSEYITGVLTVFEVSLIGAFSIGQSPSTEAFANRGAAYEIFKLIIDKPKSIDSKS 382
OY 419 GKLDPEEGHIEIRNKHAIYPSREPTVMDVLSMPAGKTALVGGSGSKSVGLVE 478
Db 383 GHRPDIKGNLEKKNHFSYSPSRKEKILKGLMKVSGQTVALVNGSGCKSTTVOLMO 442
OY 479 RFAPRGTVLLDGHDKDNLMLRQOISLVSEOPVLFGTITKYNRHGIGTKYNES 538
Db 443 RLVDPTDGMVCIQDODRTITNKHLEITGVSOEPVLFITTAENIRIG-----RENYT 497
OY 539 EDKVRLEIENAKMANAHDEFTALPEGETNVGCGFLSGGOKORIAARAAYSDPKIL 598
Db 498 MDE-----TEKAVEANAYDFTMLKPNKFDPVLVBGARLSGGOKORIAAALVBNPKIL 553
OY 599 LLDGATSAULTKSGCVQAALERAEGRTTIVAHRLSTIKTANIVYLVNGKTAEGT 658
Db 554 LLDGATSAULTKSGCVQAALERAEGRTTIVAHRLSTIKTANIVYLVNGKTAEGT 613
OY 659 DELVDGAGVRKLEAORINEOKFADLEADADLTNADIAKKTASASDDDGKPTT 718
Db 614 DELKKEGIFKLTVMOTRGNEID--LENNYGEKSESDLESPDSS-----L 663
OY 719 IDRTGTHKSVSALLSKRP-----PETPKYSILWTLKRVASPNBEPILPYMLIGLVS 771
Db 664 IKRSTRSRSHAPGGOKRIGTKEDLNENVPVSFWMLK-----LNSNEMPEYVVGJECFA 719
OY 772 VLAGGGPTQAVLYAKAISTLSLPESOYSKLRHDPADWMSLMEFYVGIOTFTOSTNGAAF 831
Db 720 IINGGLOPASIIFSRIGIIFTDEDEPKT-RONSNNFSVLEVLGIIISFTFPLGPTF 778
OY 832 AVCESELRIRARSTAFRTIIRADIAFDKEENSTGALTSFSTKHLSSGVTIGLIL 891
Db 779 GKAGEILTKRIRYVFSMRLQDYSWDDKRNITGALTTRILANDAAQYKGAIGSLAVIT 838
OY 892 MTSTLGAIIIALAIGMKTALVCISVVPVLLACGFYRFYMLAOPFSKSLAYEGSANA 951
Db 839 ONTANIGTGIIISLIYGMQTLTLLATVPIAIAIGVEMKMLSGOALKDKKELBAGKRIA 898
OY 952 CEATSSIRTYASTLTREDWEITYHAQIDAGRTSLISVNSLLYASSQALVFCVALGF 1011
Db 899 TEALENFRYVSLTREOKFEYMYAQSLOVPYRNSLRKRAHIFGVFSITQAMMYFSYAGCF 958
OY 1012 WYGTTLGHHEDYIFRFVFCSEILFGNOSAGTVEFSFADMGKAKNMAAERRLIEDRPQ 1071
Db 959 RFGAYLVANEBMNFQDVLVFSALVFGMAAGVYSFPADAKKAVSAHAYIMITEESPL 1018
OY 1072 IDNMSBEGKLEVEGEIEFRNVHFRYPTRBEPVLRGLDITVRPGOVALVGPSCGKS 1131
Db 1019 IDSISPGIKRNTLEGANTFNEVEYFNPTRPDIPLVGLSLEVKKGQALVYSSGCGKS 1078
OY 1132 TTIALLEFPDIALGSLVDGKDISKLNINSYRSFSLVSOEPPLVYOGTIENTILGIVE 1191
Db 1079 TVVOLLEFRYDPLAGSVLIDKEIKHLNVOMLRAHLGIVSOEPPLIFDPSIENIAYGNS 1138
OY 1192 DDVEEELFIKACRDANIYDFISLPEGFNTVVGSGKMLSGOKORVATIALALDPKIL 1251
Db 1139 RYVSHHEELMQAAKANINHHFETLEPKYNTNVGDKGTQLSSGOKORIAAALVBNPKIL 1198
OY 1252 LLDGATSAULDESEKVVQAALDAARGTTIVAHRLSTIKADVIYVFDGKIVESGTH 1311
Db 1199 LLDGATSAULDESEKVVQAALDAARGTTIVAHRLSTIKADVIYVFDGKIVESGTH 1258
OY 1312 SELVQKGRITVELVNLQSLK 1332

```

```

Db 1259 QQLAQKGIYFMSISVOAGAK 1279
RESULT 12
ID 002793 PRELIMINARY; PRT; 1285 AA.
AC 002793;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Multidrug resistance protein-1.
GN MOR1.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_Taxid=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Longley M., Crawford A.M.;
RT "Ovine mdrl gene."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; U78609; AAB58489.1; -.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR001140; ABCtransport.
DR InterPro; IPR003439; ABC_transport.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transport; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR AP; binding; Transport.
KW AP-binding; Transport.
SQ
SEQUENCE 1285 AA; 142020 MW; 90153617C44856F CRC64;

Query Match 36.98; Score 2481.5; DB 6; Length 1285;
Best Local Similarity 41.88; Pred. No. 1.2e-137;
Matches 535; Conservative 230; Mismatches 467; Indels 49; Gaps 7;

OY 81 LPEDEROVL--KTULEEKVNIISFGMLRWYATKMDILIMVSTICAI----- 125
Db 17 LKROKRRFSKDEKKEKRPVSTFTMFRYSNMDRLCLMVGITLAIITHGAGLPLMLVYF 76
OY 126 -----AAASTFORIMLYQISYDEYD-----ELTKNLYEVYLGIGEFYVYV 168
Db 77 GDMTDSFAGAGNLGNTLSNMTSTIDRTREYKGLKEKEMTYAYYSGIGAGVLIAYI 136
OY 169 STVGFIYGEHATQKIREYIESILKQNIYGFDKLAGEVTTTITADTNLIQDGISERV 228
Db 137 QVSEFICLAGROVHRIRKQFPHAIMQDEIGWPDVHDVCKLMTRLNDVSKINEGIGDKIG 196
OY 229 LTLTALFVAFITAYVYKMKIALICSTTVALVLMGGSQFIIKYSKSIDSGAG 288
Db 197 MFPQAMATFFGFIPTGTGNMLTVIATLIPVGLSALYAKKIIISFTDKKLAYAAAG 256
OY 289 TVAEVSISSINNAATFGTODKIAKQYEVHLDDEAEKMGTKNOIYVGMIGAMFGIMSNYNG 348
Db 257 AVAEVYLAATVYALVFGQKKELEKRYKNKELEAKRIGIKAITANISIGAAFLITAYSA 316
OY 349 LGFMGSRFLVDGAVDGIITLVMAILLIGSFISGVNPAQFTNAVAAAKIFGTIDR 408
Db 317 LAFWGTSLVLSREYSIQOVLTVFESVLIGAFSIGQASPNTEAFANRGAAYEYFKIIDN 376
OY 409 QSPLDPSYNEGTLDHPEGHIEIRNKHAIYPSREPTVMDVLSMPAGKTALVGPSCGS 468
Db 377 KPSIDISYNTGKRPKNIGNELEFRNVHFRYPTRBEPVLRGLDITVRPGOVALVGPSCGS 436
OY 469 GKSTVGLVERFYMPCVTGLDGHDKDNLMLRQOISLVSEOPVLFGTITKYNRHGIGTKYNES 528
Db 437 GKSTVGLVQRLYDPTBEGWVSIDGODRTITNKHLEITGVSOEPVLFITTAENIRIG 496
OY 529 LGITYENESDEKVELIENAKMANAHDEFTALPEGETNVGCGFLSGGOKORIAA 588

```



```

Db 497 -----REDVTMEIOKAVKEANAYDFIMLPKNEFDLVBGRKAQSLSGOKORATA 547
OY 589 RAVSDPKILLDEATSAIDTKEGVQAALERAABRTTIVIAHRLSTIKTANIVLV 648
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 548 RALVRNPKILLDEATSAIDTKEGVQAALERAABRTTIVIAHRLSTIKTANIVLV 607
OY 649 NGKIAOGTHDELVDGAGARKIVEQORINEQKADLEADNADLNADIAKTRASSA 708
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 608 DGATVEGSHDELMDGRKIGYFKLVMTQKGNLE---LENTPEESLSKTDLMSSQDSR 664
OY 709 SSDLDKPTTIDRTGTHKSVSSALISKRPEPTPKYSIMTLTKFVASFNPEIPYMLIGL 768
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 665 SSILRRKSTRSRIRSGOSODRKSTBETDESPVPSFKRLK-----LNTPEYFVGV 720
OY 769 VRSVLAGGQPTQAVLYANAISFLSPESQYSLKRDADFWSLMFVVGIIQETSTONG 828
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 FCALINGALQAPASVIFSRIGIFTRNDNETK-RONSNFLSLLFILIGIISITFEFLQ 779
OY 829 AAFVASEELIRARSTAFRTIKODIAFPDKENSTGALTSFLSTETKHLSGVYTLG 888
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 780 FFFGKAGELITRRLRYLVRSMRLQDVSWFDDKNTGALTIRLANDAQAQVAGVSRILA 839
OY 889 TILMTSTTGAALIIAIIAGMKLALVCISVVPVLACGFYRFYMLAOFGSRKLAVEGSA 948
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 840 VITQNIANIGTGIISLIIGMOLTLLLAIVPIIAVGIEMKMLSGALKKKELEGAG 899
OY 949 NFACETASSIRVYASLTREDDWEIYHAOLDAGRTSLISVRSLLYASSQALVEFCVA 1008
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 900 KATEBAIENRFRVSLTREEREPEYMAOSLOVYRNSIRAHVFGITFSITQAMMFYSVA 959
OY 1009 LGFWYGGTLGHHEYDIFFEVCFSEILGSAQAGVFSFAPDMGAKNAAFRLFLDR 1068
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 960 GCFRRFAYLYAOGIMEFOVLLVFSANVFGAAMAGOVSTFADYAKAKVSAAVNITIKR 1019
OY 1069 KPOIDNMSBEGRKLEVEGEIEFRNHFYRPTREPOVLKGLDLYVPCQYALVPGSC 1128
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1020 IFLDYSSTEGKLPSTVESVAFNDVFNYPTRPDVPLRGSLSEVKKQOTLALVSSGC 1079
OY 1129 GKSTTALLERYDAIGSLYVDGDISKLTNINSTRYSFLSVSQEPTLYOGTITENILG 1188
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1080 GKSTVQVLLERYEYDGLVGFIDGKEVKQLNVMIAHNGIYSQEPILDCSIGENIAYG 1139
OY 1189 IYEDVPEEFLKACDANIYDFIMSLPEGFMTVYVSGKGMISGGOKOVATARALLRDP 1248
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1140 DMSRVYSGEIEHAKEANIHSTFEMLPKTYNRVGDGKTQSLSGOKORIALARALVROP 1199
OY 1249 KILLDEATSAIDSEKVVQAALAAARGRTTIAVHRLSTLOKADVIYVEQGRIVES 1308
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1200 HILLDEATSAIDSEKVVQAALDRAKRGRTCIYIAHRLSTIQNDLIVYQNGRIKEH 1259
OY 1309 GTHSELVQKGRYELVNIQS 1329
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1260 GTHOOLAKQGIYFTWVSVOA 1280

```

RESULT 13

```

ID 060502 PRELIMINARY: PRT: 1169 AA.
AC 060502:
DT 01-NOV-1996 (TREMBlrel. 01. Created)
DT 01-NOV-1996 (TREMBlrel. 01. Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21. Last annotation update)
DE P-glycoprotein.
GN PGP-1.
OS Cricetus sp.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LONG;
RX MEDLINE=91154265; PubMed=1671863;

```

```

RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced ppg-1 transcripts in multidrug
RT resistant Chinese hamster lung cells.";
RL J. Biol. Chem. 266:4545-4555(1991).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: M59234; AAA37005.1; -.
DR InterPro: IPR003593; AAA_Arase.
DR InterPro: IPR001140; ABCtransport.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00664; ABC_membrane. 2.
DR Pfam: PF00005; ABC_tran. 2.
DR ProDom: PD000006; ABC_transportr. 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ
SEQUENCE 1169 AA; 128939 MW; 72E25B7CE29DC185 CRC64;

```

Query Match 36.2%; Score 2435.5; DB 11; Length 1169;

Best Local Similarity 43.0%; Pred. No. 5.3e-135;

Matches 509; Conservative 217; Mismatches 442; Indels 17; Gaps 5;

```

OY 148 LTKNVLVYVYLGIFGVYVSVVGFYVIGEHATKQIREYVLESILRONIGYDKLAGE 207
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MTTVAYTYTGIGAVTLVATVOYFWCLAGROIHKIRKFFHAINNOETGWEDVADVE 60
OY 208 VTRITADTNLIDGISEKVGTLTALATFVAIIAYVYKMLICSTIYALVLTWG 267
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LNTRLDQVSKINEGIDKIGMFQAMATFFGGFIIGTRGMKTLIVIIAISVLSLQ 120
OY 268 GGSQFIKSKSLSDYTGAGGYVAEYVSSIRNATAFGTQDLAKQEVHLEAEKRWGR 327
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 IMAKILSFTDKELQAVANAGAAVEVLAIRVIAFGQKKELEYNNNLEAKRLGIR 180
OY 328 NOIVMGFMGAMFGVSNVNGLGFMGSRFLVDAVDGDIIVYMAIILGSLGNVSR 387
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 KATTANISKGAFLIYASTALATFYTGLSVLSKEYSIGVLLVFPVLAIPSLQASBP 240
OY 388 NQAFTNVAAMAAKIFGTIDROSPLDPSNEGKTLDFEGBIELRWVKHIYPSRPEVYV 447
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 NIEAFANAGAAVEIENIDNKRPSIDSFNGKYGKPNINGNLEFNKIHHSYPSRKVOYL 300
OY 448 EDVSLSMRPGKTYALVGPSSGSKSTVYGLVERFYMVRGCTVLGDHIDILNRLWROQI 507
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 KGLINVOGGQVVALVAGNSCGSGSTVOLLQRLDYPTBEVGSIDGODITINRYRETI 360
OY 508 SLVSGEPVLFGTIYKINRHGLIGTKYENESKEVKELEENAKMANANDFTALPEGE 567
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 GYVSGEPVLFATTIENIRYG-----RENVYDDE-----IEKAVKEANAYDFIMKLPKFD 411
OY 568 TNVGRGFLLSGGOKORIALARAVSDPKILLDEATSAIDTKEGVQAALERAABRGRT 627
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 412 TVYGERGAQLSGGOKORIALARALVRNPKILLDEATSAIDTKEGVQAALDRAKREGRT 471
OY 628 TVIAHRLSTITAHNIYVLVNGKIAEGCTHDELVDGAGARKIVEQORINEQKADLE 687
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 472 TVIAHRLSTVNRADIIAGFDGVIVQGNHELMENKGIYFKLVMTQYAGNEIGNEY 531
OY 688 DADAEDLTNAODIAKIKTASSASSDLDGKPTTIDRTGTHKSVSSALISKRPEPTPKYSIM 747
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 532 GSKNEIDMLDMS---SKSASSLIRKSTRSRIRKPHQDRKLSTKELDDEVPISRW 588
OY 748 TLLKVFASENRPPIYMLIGLVFSVLAGGQPTQAVLYANAISFLSPESQYSLKRDAD 807
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 589 RLTK-----LNSSEWYFVAVGICFALVNGALQAFATIFSKVGVYTRNDDDETK-RHSDN 643
OY 808 FWSLMEFVVGIIQETSTONGAFAVCSERLIRARSTAFRTILKODIAFPDKENSTGCA 867
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 644 LFSLLFLIGVISTFTFFLQGFTEGKAGIILTKRLRYMVFESMLRQDVSWFDPNKTGGA 703
OY 868 LNSFLSTETKHLSGVSVGLTGIIMTSTLGAALIIAIIAGMKLALVCISVVPVLACGF 927
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 704 LTRLANDAGOVKATGARLAVITQNIANIGTGIISLIYGMOLTLLLAIVPIIAIGV 763

```

QY 928 YRFNMLAOFOSRKLATVEGSAFNCATSSIRIVASLITREDVWEIYHAQIDAGCRISLI 987
 DB 764 VEMKMLGOALDKRKELEGSGKITEAIENFRIVSLTRECKFENMAYQSIQIPRYNRLK 823
 QY 988 SVLSSLLYASSQALVEFCVALGFWGGTLLGHHEYDIFRFECVSESLFGAQSAGTVES 1047
 DB 824 KAHVFGITFSTQAMKTFSTACRFGAYIVARELMPFENVLVSALVFAGMAAGOVSS 883
 QY 1048 FAPDMKAKNAFAEFRRLFRKPOIDMWSEGEKLEVEGEIEFRNHFRRPTREPCQVYL 1107
 DB 884 FAPYAKAKYASAHIIIEKVPISIDYSTGGLKPNLTGKVFKEVFNPTREDIYVL 943
 QY 1108 RGLDLTVKPGQYVALVPGSGCKSTTIALERFDALAGSLVDGKDISKININSYRSFL 1167
 DB 944 OGLMEKAKGGTLLVSSSGCKSTVVOILLERFYDPAGATVFLDGKEVNOALVOMLRAHL 1003
 QY 1168 SLVSEPTLYOGTIEKENILGLIVEDVPEEFLIRACRKNANTYDFMSLPEGFNTVYSGK 1227
 DB 1004 GIVSOEPLFCSTAEINATYDNGSRVVSQDEIERAKENAHOFIESLPDKNTFRVGDKG 1063
 QY 1228 GMLSGGQKORVAIRALRLDRPKILLDEATSAIDSESEKVVQALDAARGTITIAVHR 1287
 DB 1064 TOLSGGQKORAIARALVRQHIILLDEATSAIDSESEKVVQALDKARREGRTCIIVIAHR 1123
 QY 1288 LSTIQKADVIYVPOGKIVESGTHSELVQKRGRIYELVNLOSCK 1332
 DB 1124 LSTIQADLLIVIONGKVEGHTQOILLAKIGIYFSMVSVOGAK 1168

RESULT 14

Q8R427 ID PRELIMINARY; PRT; 1275 AA.

AC Q8R427; 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ATP-binding cassette protein 81b.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
 RA Yabuuchi H., Ishikawa T.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY082609; AAL92458.1; -
 KW ATP-binding
 SQ SEQUENCE 1275 AA; 141222 MW; 7000E5F6A0E0E19D CRC64;

Query Match 36.0%; Score 2417.5; DB 11; Length 1275;
 Best Local Similarity 40.8%; Pred. No. 7e-134;
 Matches 517; Conservative 231; Mismatches 461; Indels 57; Gaps 8;

QY 94 EIKVNIISFPGIMKRYATKMDLILIVISTICAIASAATFORIML----- 136
 DB 26 KEKRPAGVIGFMYADWLKICMAIGTIAIHHGTLLPLMLVGYMTDSFTQATRIIL 85
 QY 137 -----YQISYDFEYDELKKNLYLFYVLIGISFVTVYVSGTIGYGEHATQKI 184
 DB 86 PSVTNQSINSTQVSDSSLEEDMAMVAYYTGIGAGVLIYAVIYQVSLMCLAGAOIHRI 145
 QY 185 REYLESILRONIGYFDKIGAGEVYTRITADTNLIQDGISEKVGTLTALATVAFATIA 244
 DB 146 ROKFPHALINQIGFVDNAGELNTRLDVSKINDIGDKLGMFPQSITTFSGFIIG 205
 QY 245 YKTKWKLALCSSTIVALVLTWGGSQPTIKYSKSLDSYGAGGVAAEEVVISINNAATF 304
 DB 206 FTSQKTLVLIVLAVPLIGLSSAMMAKVLTSFTNKELQAYAKAGAVAAEEVLAIRTVIAF 265
 QY 305 GTQDCLAOYEVLHDEAKMGTKNOIYMGFMIGAMFGLMYSYVGGFMNMSRFLVDGAND 364
 DB 266 GQOKELERYNKNLEAKRVGIIKAITANISIGIAYLIVYASYALAFWYGTSLVLSNEYS 325

QY 365 VGDILITVMAILIGSFSLGANSVPAQAFTNAVAAAKIFGTIDRQSPIDPYSNECKTLDH 424
 DB 326 IGOVLTFEFSILGTFSIGHLAPNIEAFANAGAAVEIFKIIDNPSIDSTGKHKPDS 385
 QY 425 FEQHLELNKVIHYSRREYVIMEDVSLSMRPAKTTALVPGSGSKSTVYGLVEFYMPV 484
 DB 386 IMGNLEFNKNVFNYSRSEVKILKGLNLKVKSGQVVALVNGSGCKSTVYDLOLQLYPI 445
 QY 485 RGTVLLDGHDKDLNRLMLROQISLVSOEPLYFGTTIKNIHGLIGTKYENESDCKRE 544
 DB 446 EGEVSDIDODIRITVNRKILIGVSOEPLYFATTIAENIKYG-----REMYTDE--- 497
 QY 545 LIENAKANAHDELTALPEGETNVGQGFLLSGGQKORIALARAVYSDPKILLDEAT 604
 DB 498 -IEKAVKERNAYDFIMKLPKHKENTLVGERGAQLSGGQKORIALARAVYSDPKILLDEAT 556
 QY 605 SALDTRSEGVQOALERAEGRTTIVIAHRLSTITANIVLVNKGKIAEGSTHBEYDR 664
 DB 557 SALDTRSEBAVQOALDKAREGRTTIVIAHRLSTVNRNAVYIAGFDGVTYEGDNHBEIMKE 616
 QY 665 GGAYRKLYEAO-RINE---QKADALDEDAEDLTNDIAKIKTASSASDLDGKPTTI 719
 DB 617 KGIYFKLVMTQTRGNEIEPGNNAYESQSDTGASELTSEKSKPLRSTRSRHRRDOE 676
 QY 720 DRTGTHKSVSAIISKRPETTPKYSLMTLTKFVASFPNPEIYMLIGLVESVLAGGQOP 779
 DB 677 RLTSKEDVD-----EDVPMVSFWQILK-----LNISEMPYLIVGLCAVINGCIQP 723
 QY 780 TQAVLYAKAISTLSLPESOYSKLRHDADFWSLMEFVVGIIQITTOSTNGAARAVSERTI 839
 DB 724 VFAIYFSKIVGVFSRDDHETKQR-NCNLFSLFLVGMGISVYTFPOGFAGKAGEIILT 782
 QY 840 RRASTARFTILRODIAFDFKEENSGALTSLSTETKHLISGVSGVLTGTTILMTSTTGA 899
 DB 783 KLRATMVKRSMKRODISWFDHKNTGSLTTLASDASVNAKMSRLAVVQNVANLCT 842
 QY 900 AIIIALAIGMKLALVCISVVPVLLACGEYRFYMLAOFOSRKLATVEGSAFNCATSSIR 959
 DB 843 GIIISLVYGMQTLTLLVVIPLIVLGITEMKILSGQALDKRKELEISGKITEAIENFR 902
 QY 960 TVASLITREDVWEIYHAQIDAGCRISLISVLSLLYASSQALVFCVALGFWGGTLLG 1019
 DB 903 TVVSLTRQKFEFTMAQSIQIPRYNRLKKAHVFGITFACFTQAMIFYSAACRFGAYIVA 962
 QY 1020 HHEYDIFRFECVSESLFGAQSAGTVEFSPAPDMKAKNAFAEFRRLFRKPOIDMWSEEG 1079
 DB 963 RELMFPENMLVFSAVVGCMAAGNTSSPAPDYAKAKYASAHIIKRIETPIDISTSTG 1022
 QY 1080 EKLETVGEIEFRNHFRRPTREPCQVYLGLDLTVKPGQYVALVPGSGCKSTTIALER 1139
 DB 1023 LKPNMLEGNVKNGVFNFPTRPNIPVLOGLSEYVKKGTALVYSSGCKSTVVOILLER 1082
 QY 1140 FPDALAGSLVDGKDISKININSYRSFLSVQOEPLVYOGTIEKENILGLIVEDVPEEFL 1199
 DB 1083 FYNPAGVTFLDGKEIKOLNOMLRAHILGIVSOEPLFCSTITENATYDNGSRVVSHEE 1142
 QY 1200 IKACKDANTYDFMSLPEGFNTVYSGKGMLSGGQKORVAIRALRLDRPKILLDEATSA 1259
 DB 1143 VAAAREANIHOFITDSPEKYNTRVGDKGTOLSGGQKORIALARALVRQHIILLDEATSA 1202
 QY 1260 LDSESEKVVQALDAARGTITIAVHRSLSTIQKADVIYVPOGKIVESGTHSELVQKRG 1319
 DB 1203 LDTSESEKVVQALDKARREGRTCIIVIAHRLSTIONMDLIVIONGKVEGHTQOILLAKG 1262
 QY 1320 RYVELV 1325
 DB 1263 IYFSW 1268

RESULT 15

Q90235 ID PRELIMINARY; PRT; 1348 AA.

